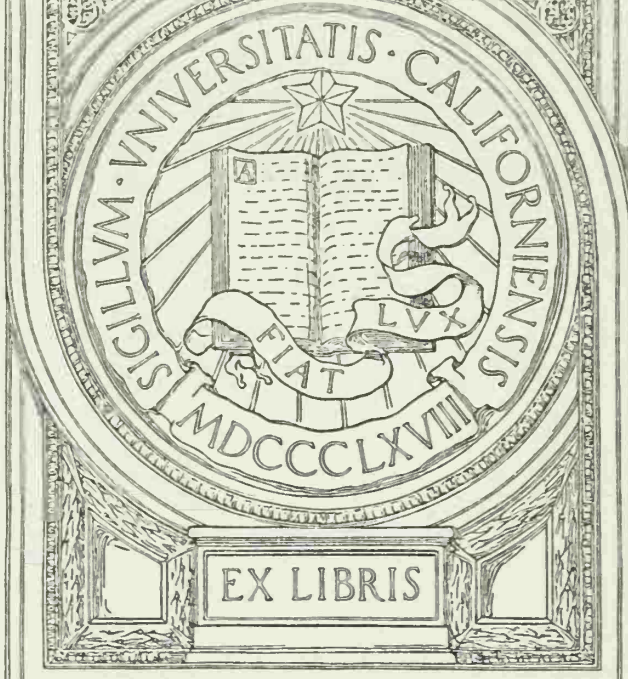
PSYCHOLOGICAL TESTS
OF
EDUCABLE CAPACITY

PRICE TWO SHILLINGS NET

MEDICAL SCHOOL
LIBRARY



EX LIBRIS

Digitized by the Internet Archive
in 2012

http://archive.org/details/reportofconsulta00grea

Gt Brit

BOARD OF EDUCATION

Report of the Consultative Committee on Psychological Tests of Educable Capacity and their possible use in the public system of education

LONDON:
PUBLISHED BY HIS MAJESTY'S STATIONERY OFFICE
To be purchased directly from H.M. STATIONERY OFFICE at the following addresses:
Adastral House, Kingsway, London, W.C.2; 28, Abingdon Street, London, S.W.1;
York Street, Manchester; 1, St. Andrew's Crescent, Cardiff;
or 120, George Street, Edinburgh;
or through any Bookseller.

1924.

Price 2*s.* 0*d. Net.*

ROYAL DENTAL HOSPITAL SCHOOL

PREFACE.

It will be remembered that when the Consultative Committee was reconstituted in July, 1920, the Board gave them two references. The first of these, upon the differentiation in curriculum between boys and girls in Secondary Schools, was discharged late in 1922, and the Report of the Committee has been for some time before the public. The second reference was in the following terms :—

> "What use can be made in the public system of education of psychological tests of educable capacity ?"

The Committee have now presented their Report upon this subject to the Board and the Board are arranging for its immediate publication.

In so doing they must not, of course, be understood as committing themselves to endorsement of the conclusions of the Committee or of the views stated by their officers in evidence before the Committee. They believe, however, that the full and thorough treatment accorded to the subject in this Report will afford valuable guidance to all those students of education—and there are many—who are anxious to pursue a topic, comparatively speaking so new, so intricate and so full of possibilities of error, yet so important ; and they desire, in the name of the general body of teachers and administrators, to thank the Committee for the time and labour expended upon the work, including of course those distinguished men who, though not members of the Committee, agreed to serve on Dr. Adami's Sub-Committee, and who in that capacity rendered indispensable assistance in the preparation of the Report.

L. A. SELBY BIGGE.

June, 1924.

LB1131
G48
1924

TABLE OF CONTENTS.

	PAGE
Preface	ii
Names of the Consultative Committee	iv
Terms of Reference	iv
Analysis of Report	v
Introduction	xi
The Committee's Report	1

Appendices :—

Appendix I.—List of Witnesses and of Persons who sent Memoranda to the Committee	146
Appendix II.—Short Accounts of some experiments recently conducted in England in the use of Group Tests and Individual Tests in Free Place Examinations and in Schools of different types	149
Appendix III.—Notes by the Secretary on the use of Psychological Tests of various types in Foreign Countries	164
Appendix IV.—Note by Dr. Cyril Burt on Standardisation and Norms	183
Appendix V.—Note by Dr. Cyril Burt on Correlation as applied to mental testing	185
Appendix VI.—Note by Mr. A. E. Twentyman on Grades in American Schools	190
Appendix VII.—A short list of recent publications on Psychology and Psychological tests	191
Appendix VIII.—Examples of tests	199
Appendix IX.—The views of various psychologists on— (*a*) The factors involved in "general" ability, and (*b*) The need for testing for "special" and "group" abilities	225
Index	239

Note.

The estimated gross cost of the preparation of the appended Report (including the expenses of the Witnesses and Members of the Committee) is £927 13*s.* 0*d.*, of which £190 10*s.* 0*d.* represents the gross cost of printing and publishing this Report.

(13579) Wt. 27584/4499A/1702 2500 7/24 Harrow G.31
(16042) Wt. 15560/1045 3000 11/24 Harrow G.31

49882

NAMES OF THE MEMBERS OF THE CONSULTATIVE COMMITTEE.

SIR W. H. HADOW, C.B.E. (*Chairman*).
MR. P. W. H. ABBOTT.
DR. J. G. ADAMI, C.B.E.
MR. S. O. ANDREW.
DR. ERNEST BARKER.
MISS E. R. CONWAY.
REV. DR. D. H. S. CRANAGE.
THE RT. HON. LORD GORELL, C.B.E., M.C.
MR. IVOR H. GWYNNE.
MISS FREDA HAWTREY.
MR. P. R. JACKSON.
SIR STANLEY M. LEATHES, K.C.B.
MR. A. J. MUNDELLA.
DR. BERTHA S. PHILLPOTTS, O.B.E.
DR. R. H. PICKARD.
MR. FRANK ROSCOE.
DR. R. P. SCOTT.
MISS. E. M. TANNER.
MR. R. H. TAWNEY.
MR. W. W. VAUGHAN, M.V.O.
MR. J. A. WHITE.

MR. R. F. YOUNG (*Secretary*).

TERMS OF REFERENCE.

What use can be made in the public system of education of psychological tests of educable capacity?

ANALYSIS OF THE CONSULTATIVE COMMITTEE'S REPORT.

CHAPTER I.—HISTORICAL SKETCH OF THE DEVELOPMENT OF PSYCHOLOGICAL TESTS.

SECTION		PAGE
1.	I.—The Development of Experimental and Individual Psychology	1
2.	II.—Physical Methods of Diagnosis	2
3.	Anatomical stigmata	2
4.	III.—The Origin of Mental Testing	4
5.	Early experimental work	4
6.	IV.—The Use of Simple Sensory Tests	6
7.	Lower Senses : Tests of touch and muscle-sense	7
8.	Higher Senses : Tests of hearing and vision ..	8
9.	V.—The Use of Simple Motor Tests	9
10.	VI.—General Results of Earlier Methods ..	11
11.	Experiments on formal training	11
12.	Negative conclusions	12
13.	VII.—The Application of Correlational Methods ..	13
14.	The coefficient of correlation	13
15.	The conception of general ability	14
16.	VIII.—Tests of Higher Mental Processes	16
17.	The possibility of group testing	19
18.	Specific abilities	19
19.	Mental imagery	19
20.	IX.—Individual Tests of Intelligence : The Binet-Simon Scale	22
21.	The diagnosis of mental deficiency	23
22.	Simplified methods	23
23.	Intelligence measured by an age-scale ..	27
24.	Revisions of the Binet-Simon Scale	29
25.	(i) The Vineland Revision	30
26.	(ii) The Yerkes Point Scale	30
27.	(iii) The Stanford Revision	31
28.	(iv) The London Revision	32
29.	(v) The Treves-Saffiotti Method	33
30.	The De Sanctis Tests	34

CHAPTER I.—*continued.*

SECTION		PAGE
31.	X.—GROUP TESTS	35
32.	XI.—PERFORMANCE TESTS	40
33.	XII.—STANDARDISED TESTS OF SCHOLASTIC ATTAINMENT..	43
34.	The Standards of former Codes of the Education Department	43
35.	American tests of scholastic attainments ..	44
36.	English tests of scholastic attainments.. ..	46
37.	XIII.—TESTS OF VOCATIONAL APTITUDE	47
38.	(i) Vocational guidance	48
39.	(ii) Vocational selection	49
40.	Vocational testing during the war	50
41.	Vocational guidance among English school children	51
42.	Vocational selection for Trade Schools and Apprentice Schools	52
43.	XIV.—TESTS OF TEMPERAMENT AND CHARACTER	53
44.	The testing of neurotic and delinquent children	53
45.	The influence of emotional and moral factors on the testing of intelligence	54
46.	(i) Tests of temperament and emotion ..	54
47.	(ii) Tests of character and morality	58
48.	The importance of observational methods as distinguished from experimental	59
49.	XV.—CONCLUSION	60

CHAPTER II.—GENERAL SUMMARY OF THE AVAILABLE EVIDENCE BEARING ON THE PROBLEMS CONNECTED WITH THE VARIOUS TYPES OF PSYCHOLOGICAL TESTS OF EDUCABLE CAPACITY.

SECTION		PAGE
50.	The various types of psychological tests of educable capacity	62
51.	Observed discrepancies between ability and attainment; the desirability of recourse to some means of discovering innate ability apart from examinations of the conventional type	63

CHAPTER II.—*continued.*

SECTION		PAGE
52.	The application of "Intelligence" Tests to subnormal and supernormal children	64
53.	What is "Intelligence" as understood by modern psychologists	67
54.	The various hypotheses regarding the nature of the factors involved in "general ability" or general intelligence	68
55.	What do Tests of "Intelligence" measure	71
56.	The main presuppositions underlying the use of Tests of "Intelligence," and the indispensable conditions which must be complied with in order to ensure the validity of any set of Tests of "Intelligence" for diagnostic purposes	74
57.	The Binet-Simon Scale: Its merits and disadvantages ..	78
58.	The opinions of medical experts on the value of the Binet-Simon Scale and its modifications as an aid in the diagnosis of mental deficiency in children	82
59.	Merits and defects attributed to the American modifications of the Binet Scale	83
60.	Advantages and defects attributed to other individual tests	86
61.	General character of group tests of intelligence and their value for determining educable capacity	87
62.	The relative merits and disadvantages of group tests and individual tests of intelligence	90
63.	The place of the interview in the application of individual tests, and its technique	92
64.	How far do tests of "intelligence" throw light on character or temperament, and how far are they affected by them. The possible use of temperamental tests as ancillary to intelligence tests	93
65.	The connection between emotion and general intelligence	95
66.	The value of standardised scholastic tests based on age performance for gauging educable capacity	96
67.	The value of vocational tests (including tests of manual ability) in determining educable capacity	98
68.	The nature and value of physical tests as supplementary to the preceding	100
69.	The danger of special preparation or "coaching" for the tests	103
70.	Suggestions by witnesses in regard to the training of those applying the tests	105
71.	Recommendations by witnesses on the desirability of establishing a central organisation to try and direct new tests and to collate experience	108

CHAPTER III.—THE VARIOUS POSSIBLE APPLICATIONS OF PSYCHOLOGICAL TESTS OF EDUCABLE CAPACITY IN THE PUBLIC SYSTEM OF EDUCATION.

SECTION		PAGE
	PART I.—The Possible Use of "Intelligence" Tests in Connection with Ordinary Examinations Intended Primarily to Test Ability.	
72.	Introduction	108
73.	Tests of "intelligence" may be regarded as a species of the ordinary written and oral examinations, in so far as such examinations are designed to test ability rather than attainments	109
74.	The First and Second School Examinations, being examinations intended primarily to test attainments, are outside the scope of our inquiry	111
75.	On the other hand, examinations for Free Places in Secondary Schools, for entrance to Secondary Schools, and for admission to Central Schools, should be primarily designed to discover ability rather than to test attainments	111
76.	A brief account of the more important expedients adopted up to the present time by various Local Education Authorities to render the examination for Free Places a more effective test of ability	113
77.	Summary of the evidence submitted to us relating to certain shortcomings in examinations for Free Places as at present conducted, and regarding the possible use of "Intelligence" Tests as adjuncts to such examinations	115
78.	Suggestions for the experimental use of "Intelligence" Tests in connection with examinations for Free Places	118
	(*a*) In the Elementary School before the Free Place examination, or in association with the preliminary qualifying or "weeding out" examination in areas where such preliminary examination is held	119
	(*b*) The use of group tests of intelligence in the examination for Free Places in areas where, even after the "weeding out" examination, a large number of candidates sit for the Free Place examination proper	119
	(*c*) The use of individual Tests of "Intelligence" at the oral interview held in connection with the Free Place examination	120
79.	The importance of the oral interview in examinations intended primarily to discover ability	120
80.	The special importance of the oral interview in public examinations for young children, such as Free Place examinations, and examinations for admission to Central Schools	121

CHAPTER III.—*continued.*

SECTION		PAGE
	PART. II.—THE POSSIBLE USE OF "INTELLIGENCE" TESTS WITHIN SCHOOLS OF DIFFERENT TYPES.	
81.	Introduction.	124
82.	In Elementary Schools:—	
	(*a*) Application of individual tests to pupils of about seven years of age on admission to the School ..	125
	(*b*) Application of individual tests to retarded children and children suspected of mental deficiency ..	126
	(*c*) Application of group tests to pupils of about eleven years of age before presentation for the Free Place examination	126
	(*d*) Application of group tests, followed in difficult cases by individual tests, to doubtful and border line pupils, about whose ability there is considerable divergence of opinion on the part of different teachers	126
	(*e*) The use of group tests, in association with the ordinary terminal examinations in Elementary Schools, does not seem advisable or necessary save in exceptional cases	126
83.	In Junior Technical Schools:—	
	Application of group tests to entrants in order to help the school authorities in classifying them	127
84.	In Central Schools:—	
	Application of group tests, followed in difficult cases by individual tests, to pupils who are considered suitable for transfer to Secondary Schools at the age of about fourteen	127
85.	In Secondary Schools:—	
	(*a*) application of group tests to pupils newly admitted to Secondary Schools with a view to their proper classification	127
	(*b*) Application of group tests, followed in difficult cases, if necessary, by individual tests, to assist the school authorities in assigning pupils, after one or two years at the Secondary School, to different divisions of a form	128
	(*c*) Application of group tests, followed up in difficult cases by individual tests, to pupils whose general ability, as revealed in their attainments, is variously estimated by different specialist teachers	128
	PART III.—THE POSSIBLE USE OF STANDARDISED SCHOLASTIC (EDUCATIONAL) TESTS, VOCATIONAL TESTS, PHYSICAL TESTS, TESTS OF MEMORY, PERCEPTION AND ATTENTION, AND TEMPERAMENTAL TESTS, AS AFFORDING DATA BEARING ON EDUCABLE CAPACITY, SUBSIDIARY TO THE INFORMATION YIELDED BY TESTS OF "INTELLIGENCE."	
86.	The possible use of standardised scholastic (educational) tests as an ancillary means of gauging the educable capacity of pupils in schools of various types	129

CHAPTER III.—*continued*

SECTION		PAGE
87.	The possible use of vocational tests, including tests of manual ability, in schools and other educational institutions	130
88.	The possible use of physical tests as a subsidiary means of assessing educable capacity	131
89.	The possible use of tests of memory, perception, attention, imagery, and association, as ancillary to Tests of "Intelligence"	131
90.	The possible use of temperamental tests as subsidiary to "Intelligence" Tests in assessing educable capacity	132

PART IV.—THE PROVISION OF COURSES FOR TEACHERS IN EXPERIMENTAL PSYCHOLOGY, WITH SPECIAL REFERENCE TO THE THEORY AND PRACTICE OF PSYCHOLOGICAL TESTS, AND THE QUESTION OF TAKING STEPS TO CO-ORDINATE AND ORGANISE RESEARCH AND TO DISTRIBUTE INFORMATION BEARING ON THE APPLICATION OF PSYCHOLOGICAL AND STATISTICAL METHODS TO EDUCATION.

91.	The question of the training which should be given to Teachers, School Doctors, and other persons engaged in the work of education to render them competent to administer group and individual tests of intelligence ..	133
92.	The question of provision for the co-ordination and correlation of the work of psychologists, teachers and statisticians in regard to the application of psychological and statistical methods to education	134

PART V.

93.	Conclusions and Recommendations	135

INTRODUCTION.

On 29th July, 1920, the Board sent two References to the Consultative Committee, one on the differentiation of curricula, the other on the use of psychological tests in the general system of education throughout the country. For reasons which will be readily appreciated, the Committee took first the Reference on differentiation of curricula, spent about two years in its investigation, and reported in September, 1922. Meanwhile, realising that its second Reference was highly technical and would need much expert assistance and advice, it appointed a Sub-Committee of six of its members, with Dr. Adami as Chairman, and with power, subject to the approval of the President, to co-opt members from outside.[1] In this way it was fortunate enough to secure the services of Dr. P. B. Ballard, Dr. C. S. Myers, and Professor C. E. Spearman, who placed at its disposal their wide knowledge and sound judgment and to whose ungrudging help and co-operation it is very deeply indebted. Not less does it owe its cordial gratitude to Dr. Cyril Burt, who supplied it with evidence, who wrote for it some invaluable memoranda for Appendices IV, V, and VII, and whose work has influenced almost every page of our first chapter.

The Sub-Committee sat 13 times between October, 1920, and May, 1922, examined 19 witnesses, and received in addition many valuable pieces of written evidence from psychologists at home and abroad. On 25th May, 1922, it presented to the full Committee a report in which the evidence before it was digested and interpreted. This report the full Committee took as the basis of its work.

Since May, 1922, the full Committee has sat on 22 days, and has heard evidence from 18 witnesses, including the three experts who had served on the Sub-Committee and who very kindly came to explain and develop certain of the points raised in its Report. A number of additional memoranda, scientific, educational and statistical, have also been received and have been much utilised especially for the Appendices. Many members of the Committee have visited schools and institutions where the tests were in use in order to investigate at first hand the methods by which these are applied, marked and interpreted. They also visited the laboratories of Dr. Myers and Professor Spearman, and submitted themselves to experiment.

---

[1] See Clause 5 (ii) of the Order in Council, 22nd July, 1920.

From the outset the Committee decided to interpret the term "test of educable capacity," in its widest denotation (*see* Report, § 50); and after consideration of the evidence from all sides, came to the conclusion that only three types of tests, the so-called intelligence tests, standardised scholastic tests, and in a less degree, vocational tests were suitable for use in schools and other educational institutions as distinct from psychological laboratories. Even within these limits it appeared that the theory of mental testing has not advanced as yet beyond the stage of what may be called temporary stability. It would therefore be dangerous and misleading to lay down any dogmatic rules as to the uses and limitations of the various types, especially as, up to the present time, comparatively few systematic experiments in their use have been made through England and Wales, and there is consequently no large body of experience suitable to national conditions such as is available in the United States and in Germany. A *resumé* of the work that is being done in countries outside Great Britain is given in Appendix III.

We take this opportunity of thanking our witnesses for their evidence, much of which is very technical and must have taken great time and trouble to prepare. We regret that it is impossible to print the evidence in full, but we reproduce in Appendix IX the views of psychologists on the difficult question of testing for general and special abilities. Especially we would record our gratitude not only to the four distinguished psychologists whose names we have already cited, but to Mr. A. E. Twentyman, Librarian of the Board of Education, who prepared for us the valuable note on Grades in American Schools printed in Appendix VI, and to our Secretary, Mr. R. F. Young, whose unstinted labour has greatly helped us to thread the mazes of an abstruse and complicated problem.

CHAPTER I.

HISTORICAL SKETCH OF THE DEVELOPMENT OF PSYCHOLOGICAL TESTS.[1]

I.—THE DEVELOPMENT OF EXPERIMENTAL AND INDIVIDUAL PSYCHOLOGY.

1. For over two thousand years, in its general problems and accepted principles, psychology presented no great change or development. It continued, as it had begun, a branch of metaphysics rather than of science. The chief method of the psychologist was still introspection; and his chief subject, himself. All that he could offer to the teacher was an *a priori* system of generalised maxims, vague, speculative, commonplace, and of little practical use.

But during the early part of the nineteenth century, influenced chiefly by the introduction from other natural sciences of an experimental procedure and of a mathematical technique, the study of the mind took a new direction. The psychologist left his armchair for the laboratory: and his inquiries and his methods moved further away from those of the philosopher, and inclined for a while towards those of the physiologist. His science, however, for more than another fifty years, still remained an abstract discipline, a science pure and unapplied, pre-occupied with fundamental issues as to the nature and the working of mind in general. Differences between one mind and another it ignored. From time to time, indeed, when studying some sample person in his laboratory, the notice of the experimenter was inevitably diverted from a general analysis of thought in the abstract to the concrete peculiarities in the thoughts of particular thinkers. Yet for long these individual differences were looked upon merely as disturbing irregularities, deviations which had to be allowed for, or averaged away, before any sound conclusion of universal scope could be examined or deduced. "Personal equations," therefore, were devised to discount these so-called errors of measurement. At length, however, attention came to be focussed directly upon these differentiating qualities, in and for themselves. Their origin was investigated; and their variations with varying age, sex, race, heredity and

[1] This historical introduction was prepared at the request of the Committee by Dr. Cyril Burt.

environment were themselves discovered to be rich in interest. And thus at last an independent branch of mental science, "differential" or "individual psychology," was founded and named[1].

II.—PHYSICAL METHODS OF DIAGNOSIS.

2. In previous centuries, it is true, some crude and curious attempts had been made upon empirical lines to draw up brief rules for the estimation of individual capacity and character. Most of these efforts, however, relied upon external and physical signs, a tendency which to-day, though obsolescent, is not extinct. In the eighteenth century, Lavater, bringing together many traditional precepts, and adding several shrewd observations of his own, published a treatise on physiognomy; a man's disposition and abilities were to be inferred, according to ingenious principles, from the features and expression of his face[2]. Thirty years later Gall propounded his system of phrenology, localising some twenty-six complex propensities in various areas or "organs" of the brain, and deducing their development from the relative prominence of the overlying regions of the skull[3].

3. *Anatomical Stigmata.* Of all such attempts at inferring psychical qualities from physical signs the most recent and the most widespread is the doctrine of anatomical stigmata. This doctrine received its most thorough-going application in the theories of the Italian school of criminology, whose exponents,

---

[1] The problems, principles and development of individual psychology may be traced (down to the date of publication) in W. Stern's *Differentielle Psychologie* (1911), which contains a classified bibliography of over 1,500 references.

[2] The many teachers who are interested in the possibility of physiognomical diagnosis may be referred, not only to Lavater's quaint *Essay on Physiognomy destined to make man known and loved* (1772), but also to Bell's *Anatomy and Philosophy of Expression* (1806; 3rd ed. (posthumous) 1872), Darwin's *Expression of the Emotions of Man and Animals* (2nd. ed. 1889), Mantegazza's *Physiognomy and Expression* (1904), and, for recent statistical studies, *Child Study*, June 1919, "Facial Expression as an index of Mentality," and Langfeld, "Judgments of Emotions from Facial Expression," (*J. Abn. Psych.* xiii., 172; cf. also *Psych. Rev.*, xxv., 488.)

[3] Gall and Spurzheim. *Anatomie et Physiologie du Système Nerveux en Général et du Cerveau en Particulier avec des observations sur la possibilité de reconnaître plusieurs dispositions intellectuelles et morales de l'homme et des animaux par la configuration de leurs têtes.* (1810).

under the lead of Cesare Lombroso, published their observations and conclusions during the closing decades of the nineteenth century. The upholders of this doctrine claimed that it was possible to recognise defective intelligence and degenerate character from certain visible marks or malformations connected principally with the size and shape of the head. Small, misshapen, or asymmetrical skulls, low, narrow and bossed foreheads, broad, depressed or upturned noses, narrow, high or V-shaped palates, lobeless, projecting or crumpled ears, these and many similar anatomical anomalies were thought to indicate a reversion to some low and primitive type. Even at the present day there are still many teachers and medical officers who prefer to base their diagnoses of mental defect or moral degeneracy upon physical peculiarities of this kind[1].

Generally, however, little emphasis is now placed upon cranial stigmata. In destroying all practical reliance upon the inspection and measurement of heads, the work of Professor Karl Pearson and his followers has perhaps played the decisive part[2]. Pearson, after analysing data from 5,000 school children and 1,000 undergraduates, concluded that the correlations between intelligence and the dimensions of the skull, though positive, are far too small for trustworthy predictions. There are,

---

[1] Of early work done upon these lines in England, the most important was that of Dr. Francis Warner. Extensive examinations in various English cities were carried out by him, and reported to a Special Committee of the London Charity Organisation (1876) appointed to inquire into the education and care of imbeciles. Whatever may now-a-days be thought of Dr. Warner's reliance upon physical criteria such as "nerve signs" and "stigmata of degeneracy," his surveys did much to draw public attention in this country to the question of inborn differences of intelligence and to the problem of the subnormal child.

[2] "On the Relationship of Intelligence to Size and Shape of the Head." *Biometrika*, V. 1906. pp. 105–146. Galton was here the chief pioneer ("Head Growth in Students," *Nature* XXXVIII, 1886, pp. 14 and 40). Both he and Binet, (*L'Année Psychologique* VII, 1900, pp. 314–42; cf. also XVI, 1910, pp. 1–30) obtained slight but positive results. Later researches have emphasised the smallness of the correlation more than its positive character. Dr. Goring's research upon convicts in H.M. prisons also deserves mention as giving what some have considered the death-blow to the more extreme claims of Lombroso and the anthropometric school of criminology (*The English Convict: A Statistical Study*. H.M. Stationery Office, 1913.)

indeed, very few persons, whether normal or abnormal, who do not possess at least one or two so-called stigmata of degeneracy. Unless several of such stigmata are found together in one and the same individual, the indications mean little or nothing; and the conjunction of a number of them is rarely seen, even in special schools for educable defectives, or, in fact, anywhere beyond the walls of asylums for low-grade imbeciles, whither the so-called clinical types are now commonly removed. Except for these rare pathological cases, psychologists are nowadays agreed in distrusting all snapshot judgments based upon an inspection of the face and head; and, *when diagnosing mental characteristics, rely upon mental rather than upon physical criteria*[1].

III.—THE ORIGIN OF MENTAL TESTING.

4. Of the application of the new experimental methods to the direct study of the differences of individual minds, the most conspicuous result was the invention of mental tests. After many long and doubtful trials the possibility of mental testing has at length received general acceptance; and recently upon an enormous scale, its value has been demonstrated, and its uses popularised, by the wholesale psychological examination of recruits for the American Army.

5. *Early Experimental Work.* It is a widespread but erroneous notion, current both among teachers and the general public, that psychological testing is a foreign invention, a new and alien import brought to England from abroad; school testing, it is

[1] In each of these views—in the speculations of the physiognomist, in the fancies of the phrenologist, and in the doctrine of anatomical stigmata—there are, of course, certain elements of truth. Glandular disturbances, for example, affect both character and intelligence, on the one hand, and the development of the bony structures—most conspicuously, perhaps, the bony structures of the head—upon the other. Racial types, again, are characterised by slight differences in intellect and emotionality, and tend to exhibit distinctive traits in the conformation of the skull and face. Quite recently, the method of correlation has shown that often the inferences of practised observers from peculiarities of facial expression are trustworthy to a tolerably high degree, though here, indeed, it is not so much the shape of the hard features or of the bony framework that is significant, but rather the tonus and contractions of the facial muscles, the habitual expression of moods, and the passing responses to emotional stimuli (See references cited in footnote [2] p. 2 above).

supposed, has come from France, vocational testing from the United States. This, however, is unjust. The conception of the mental test, to whatever extent it may have been developed and applied of recent years in other countries, was originally put forward by an English scientist. It was Sir Francis Galton who in 1883 first announced the possibility of measuring intellectual abilities by simple laboratory tests[1]. Galton's interest, however, was subsequently diverted towards problems of anthropology and eugenics. He suggested, indeed, or elaborated, some of the chief statistical devices now in use—the method of percentiles and the method of correlation; from time to time, too, he designed special apparatus for testing the muscular, auditory, and other senses of different persons; he even commenced various investigations into the correspondence between anthropometric measurements and intellectual characteristics; and finally, in 1890, he appended to an article on *Mental Tests and Measurements*, by Professor Cattell[2], an ambitious proposal for comparing laboratory data with independent estimates of human qualities.

The work of Cattell and Galton may be claimed as the first seed which bore such prolific fruit in the numerous American researches of later years. Cattell's most distinguished pupil, Prof. E. L. Thorndike, of Columbia University, who has himself carried out and inspired more numerous researches upon educational tests than any other man, writes thus of his own master: "Cattell refined Galton's methods, and won recognition for the mental measurement of individuals as a standard division of psychology. . . . His work on mental tests was the first of a series of influential contributions made during the last decade of the nineteenth century; and was for many of us the introduction to the whole topic of individual psychology[3]."

Galton's early tests were carried out mainly in his Anthropometric Laboratory at South Kensington Museum. Cattell instituted a similar series of mental measurements for the students

---

[1] Galton, F., *Enquiries into Human Faculty and its Development* (1883).

[2] Dr. J. McK. Cattell (Professor of Psychology, first in the University of Pennsylvania and later in Columbia University) *Mind*. 1890, XV. 9, p. 380.

[3] *Col. Univ. Contr. to Phil. and Psych.*, XXII, iv., (1914), 92.

of Columbia University[1]. Both of them aimed, by the intensive application of a long programme of laboratory tests to separate individuals, at making "a systematic inventory of their mental traits."

Although rich in suggestions for later inquiries, Galton with his English collaborators, achieved along these lines no final conclusions of his own. And, during the closing years of the nineteenth century the notion of mental testing, in spite of its English origin, was taken over almost exclusively by foreign investigators—by such early workers as Oehrn and Ebbinghaus in Germany, and Boas and Gilbert in America, and, later on, Bourdon, Binet, Henri, and their collaborators in France.

These earliest foreign researches, carried out between 1890 and 1900, though numerous, proved contradictory and disappointing. Yet, even at the present time, they are by no means uninstructive. They exhibit, in a clear and concrete fashion, the main recurrent fallacies that beset all who attempt to apply psychological tests to practical ends.

Of these initial investigations nearly every one is characterised by two limitations: first, the experimenters confined themselves almost entirely to the simplest mental processes: secondly, they possessed no adequate statistical procedure for determining the value of their results.

IV.—THE USE OF SIMPLE SENSORY TESTS.

6. The tests first used for the measurement of mental capacity consisted chiefly of the traditional laboratory experiments upon simple sensory discrimination. For this there seem to have been several reasons. To begin with, the majority of the new experiments in pure psychology, experiments chiefly suggested by the methods of physiology and physics, were at that time concerned with sensory capacities—with the ear, the eye and the skin, the perception of sound, of light, of weight, and so forth. Further, theoretical psychologists both in England and abroad were still much under the influence of that British school of Philosophy, whose axiom had been—"*Nihil in intellectu quod non prius in sensu.*" The working hypothesis underlying all these early studies with sensory tests seems to have been an emphasis upon

[1] J. McK. Cattell and L. Farrand *Psych. Rev.*, (1896), 618–648 "Physical and Mental Measurements of the Students of Columbia University."

cognitive attention; it was supposed that the physiological limits of the several sense-organs were much the same in different individuals, but that the discrimination of the data supplied by those sense-organs depended upon the power of attentive analysis, a power which appeared to vary greatly from one man to another, and which many writers at this date were disposed to identify with general intellectual capacity.

7. *Lower Senses: Tests of Touch and Muscle Sense.* Galton, and many of the earliest investigators abroad, dealt principally with what may be termed the lower senses. In his enquiries into human faculty and its development, reported in a small volume with that title[1], he had noticed that mental ability appeared to be correlated with delicate sensory discrimination, as tested, for example, in the comparison of graded weights. Fellows of the Royal Society, he found, could distinguish tiny differences in heaviness with an accuracy equalled only by the practised sorter at the post-office: incidentally, he observed, what almost all psychological experience has since confirmed, that tests of capacity as distinct from attainment are, in general, but little affected by practice or experience—far less than would otherwise be supposed. Later, both Gilbert, in America[2], and Spearman, in England[3], confirmed this correlation[4]. Another lowly form of sensory discrimination, much favoured by the early psychophysicists, was the power of distinguishing two touches upon the skin, the touches being produced by the points of the compass, or by a special instrument known as the æsthesiometer. Binet in France[5], and later Schuyten in Belgium[6], both following an earlier German investigation by Wagner[7], found, or thought they found, that touch discrimination, tested in this manner, formed a practicable measure of intelligence. This notion, however, more accurate methods have not confirmed. With the lower senses, it would seem, at any rate within broad

[1] See Everyman's Library Edition, pp. 23–25 and 248–251.

[2] *Stud. Yale Psych. Lab.*, 1893, II., pp. 40 *et seq.*

[3] *Amer. Journ. Psychol.*, 1904, XV., pp. 201 *et seq.*

[4] The actual measurement of these correspondences by the precise method of correlation was only introduced by later investigators. See below, § 13.

[5] *L'Année Psych.*, 1899, VI., pp. 248 *et seq.*

[6] *Arch. de Psych.*, 1903, II., pp. 321–6, *Paedologisch Jaarbock*, VII., 1909, pp. 73–116 *et al.*

[7] *Samml. u. Abhandl. a. d. geb. d. Pad. Psych.*, I., 1896.

limits, young children are almost as sensitive as old, dull children as bright, and savages as civilised men.

8. *Higher Senses : Tests of Hearing and Vision.* With tests of the higher senses, of hearing and of sight, some positive correspondence with general intelligence has been more certainly established. Many of the earlier physical surveys carried out in schools had demonstrated that defective vision was extremely common among backward children, although, of course, among the short-sighted many able persons are, nevertheless, to be found.[1] And, measuring experimentally the power to discriminate different shades of brightness by the eye, Gilbert found a small but significant agreement between his visual tests and the intelligence of his examinees ; Spearman too, found a similar but larger agreement. One of the earliest of all researches upon the inter-relations between different mental capacities, the investigation of Oehrn[2], had introduced a supposed test of visual discrimination somewhat more complex, which has long remained a favourite. This was the cancellation test, consisting in the erasure of given letters of the alphabet from a page of printed matter. Both Oehrn himself in Germany, and Bourdon following him in France[3] found the results of this exercise tally closely with those of other intellectual tests ; and later writers, using as a criterion the independent judgment of careful observers, confirmed the presence, if not the amount, of this correspondence. Tests of hearing the earlier experimenters do not seem to have employed, though here, once again, previous medical inspections had shown that defective hearing impedes educational progress quite as noticeably as defective sight[4]. The auditory test preferred by the

[1] See, *e.g.*, the report of the extensive tests carried out by Smedley at Chicago : 46*th Annual Report of Board of Education* (Chicago), 1899. Cohn had endeavoured to demonstrate, not unsuccessfully, that myopia is essentially a disease of civilisation, rare among country children and common in the town : see his *Hygiene of the Eye*, 1886, and *Die Sehleistung von* 50,000 *Schulkindern*, Breslau, 1899.

[2] " Experimentelle Studien zur Individual-Psychologie," *Dorpater Dissertation*, 1869 : see also Kraepelin, *Psych. Arbeit.*, I., pp. 92–151.

[3] *Revue Philosophique*. XL., pp. 153 *et seq.*

[4] See, *e.g.*, Smedley, *loc. cit. sup.*, and for a summary of earlier investigations, Chrisman, " The Hearing of School-children," *Ped. Sem.*, 1893. II., p. 397.

psychologist consists usually in some form of distinguishing the pitch of musical notes. The measurements obtained from such a test agree with independent estimates of intellectual ability more nearly than those obtained from any other form of simple sensory discrimination. Spearman, for example, working at a later date, and using more precise statistical methods, found a correlation of 0·94 ; that is to say (as he puts it), " the intellectual function . . . is nine parts out of ten responsible for success in such a simple act as discrimination of pitch[1]." His observation has since been confirmed by other investigators, such as Whipple[2] who obtained an analogous correlation appreciable and positive, though not quite so high[3]. The peculiar relation between auditory discrimination and general intelligence has been tentatively explained by the dependence of the development of the higher intellectual capacities in man upon his power of speech, and of his power of speech upon his power of hearing.

These earlier experiments upon sensory capacity culminated at length in the detailed and elaborate research by Spearman, to which reference has already been made, and which will be described in greater detail on a later page. He found a positive correlation between all measurable modes of sensory discrimination ; and attributed this to a common factor which he termed " General Discrimination." Further, he inferred a similar common factor underlying the different manifestations of intelligence, terming this " general intelligence." And, finally, he concluded that " the common and essential element in Intelligence wholly coincides with the common and essential element in the Sensory Functions."

V.—THE APPLICATION OF SIMPLER MOTOR TESTS.

9. Meanwhile, both in England and abroad, a reaction had been gathering ground against what was held to be the intellectualistic and sensationalistic bias of traditional psychology.

---

[1] " General Intelligence Objectively Determined and Measured," *Amer. Journ. Psychol.*, XV, 1904, p. 285.

[2] *Manual of Mental and Physical Tests*, p. 221.

[3] Seashore reported no correlation between pitch-discrimination and intelligence ; and considered rather that the results of this test indicated musical capacity. Spearman, however, applying to Seashore's own data more careful methods of correlation obtained from them a distinct and positive coefficient.

Emphasis was being placed more and more upon the active aspects of the mind. In the United States especially, psychology was acquiring a dynamic and pragmatic cast. And it is, therefore, natural to find American investigators beginning to substitute tests of a more active character for the earlier tests of simple cognitive capacity. The measurement of sensation gave way to the measurement of movement.

In the new quantitative psychology one of the earliest forms of laboratory experiment had been the measuring of reaction time, of the quickness of motor response to a sensory stimulus. Using this form of experiment side by side with tests of sensory discrimination, Gilbert[1] discovered a far larger correspondence between intelligence and speed of reaction than he had observed between intelligence and the discrimination of sensations. Binet[2] also adopted the reaction-experiment; but found that the reaction-times harmonised poorly with intellectual capacity, whereas a test of pure movement—simple tapping in fact—showed, at any rate among young children, a closer agreement. Tapping has since been much employed by subsequent investigators; and most of them—Smedley, Bolton, Abelson, Kirkpatrick and Gilbert himself—have verified the existence of some such correlation. Once more, however, the correspondence obtained from all those simple motor tests has seldom proved to be conspicuously high. With older and brighter children it usually proved smaller than with the younger, duller, or defective children; and one or two careful investigators—such as Bagley and Whipple—have discovered no correspondence whatever. With strength of movement, as with speed of movement, the same low values have been found. The dynamometer, for example, may be conveniently used to test the power of grip. But here, as before, the connection with intelligence, though positive, is, as a rule, variable and slight[3].

---

[1] *Stud. Yale Psych. Lab.*, II., 1893, pp. 40 *et seq.*

[2] *L'Année Psychologique*, IV., 1897, pp. 64–98, "Épreuves de Vitesse chez les Jeunes Garçons."

[3] Binet et Vaschide, *L'Année Psych.*, IV., 1897, pp. 15–63. "Expérience de Force Musculaire chez les Jeunes Garçons," Smedley, 46th *Ann. Rep. Board of Educ. Chicago*, 1899. Schuyten, *L'Année Psych.*, IX., 1902, pp. 448–9.

VI.—GENERAL RESULTS OF EARLIER METHODS.

10. It will be noticed that, throughout this earlier series of researches, the experimenters were confining themselves almost exclusively for their tests to the simplest mental functions, functions that need for their analysis and measurement technical instruments and apparatus, often difficult to use, generally complicated in structure, and seldom very portable in size and weight. Such tests and such contrivances were rarely fitted for application to young children in the ordinary schools. From time to time, it is true, tests of memory, attention, association, or illusion, were added to the tests of sensation or movement. But even these were of a relatively elementary order. Looking back upon these methods, therefore, it is not surprising to discover that they contained no promise of a sure or certain indication for a quality so complex as educable capacity.

This slow suspicion, now beginning to break upon the believers in laboratory apparatus themselves, was ultimately confirmed by two or three investigations of paramount importance, which approached the general problem from a somewhat novel angle.

11. *Experiments on Formal Training.* Of these newer researches the first was mainly negative in its result. It comprised a group of experiments inspired by Professor E. L. Thorndike, an inquiry which has had far-reaching effect upon educational theory. Primarily, his problem was concerned, not with mental testing, but with mental training and its transfer—with the traditional doctrine of so-called formal discipline.

Hitherto, most investigators had been content to assume that each of their tests measured some general faculty or function, and that one test only was all that was needful to provide a complete measurement of the faculty or function in question. Crossing out letters was taken to measure perception ; repeating numbers, to measure memory ; reading syllables exposed for a brief instant, to measure attention. And these assumptions, as a rule, were left wholly unverified. Often there was no attempt to check the supposed correlation with faculties or with intelligence, by comparing the test-results with independent estimates of ordinary observers ; the experimenter was content merely to relate one test to another and announce that " memory is proportional to perception " or that " association varies

inversely with other intellectual powers." Conformably with this implicit belief, it was almost universally accepted that the repeated exercise of intelligence, of will, or of attention, in some limited direction or upon some special subject, would inevitably be followed by a general all-round improvement in the faculty involved[1].

Towards the close of the century, however, the experiments of Thorndike, Woodworth, and others showed very conclusively that there might be little or no connection between test-results for one and the same function, even when both tests and functions were designated by an identical name. Training in one mental activity exercised no appreciable influence upon other mental activities, even where these activities seemed to be most intimately akin[2]. And later investigators, pursuing this new line of thought, showed that there might be extremely low correspondence between the results of different tests all claiming to measure the same general function—memory, attention, suggestibility, or whatever it might be.

12. *Negative Conclusions.* The conflicting deductions of various investigators, working with similar tests and a similar procedure, thus seemed to indicate that mental testing could lead to no consistent or trustworthy result. This general disillusionment gained further confirmation from a summary of the most ambitious programme of enquiry carried out in the nineteenth century. In 1901 Wissler[3] published the result of ten years' work at Columbia University with an elaborate series of experimental tests, based largely upon Cattell's original scheme. The conclusions are about as negative as could be conceived. "The laboratory mental tests," say the writers, "show little inter-correlation. . . . The markings of students in college

[1] Stumpf, for example, had written "The power of mental concentration upon certain points, in whatever region acquired, will show itself effectual in all others also" (*Tonpsychologie*, p. 123); and again "Development of will power in connection with any activity is accompanied by development of will power as a whole" (*Psych. Rev.*, VI., p. 163).

[2] In passing it should be remarked that this conclusion forms one of the many replies to those who fear upon *à priori* grounds that tests claiming primarily to measure inborn capacities must be gravely disturbed by practice or experience in allied mental processes.

[3] *Psych. Rev. Mon. Supp.*, 1901.

classes correlate with themselves to a considerable degree, but not with the tests made in the laboratory." In Germany, so eminent a psychologist as Kraepelin[1] came also to an identical verdict. "At the end of these inquiries," he writes, "we cannot hide from ourselves that the results secured have fallen far short of what we had been led to anticipate from collective experiments with the simplest 'mental tests.'"

The sanguine hopes of Galton and his immediate followers seemed, then, to have come to nothing; and the notion of mental testing fell into temporary disrepute[2].

VII.—APPLICATION OF CORRELATIONAL METHODS.

13. After these discouraging pronouncements, for several years little work of any moment was done with mental tests. Suddenly, however, interest was re-awakened; a new impetus and a fresh turn were imparted to the whole inquiry by two of the most original and suggestive investigators in this field, Spearman in England, and Binet in France.

14. *The Coefficient of Correlation.* Following the suggestions and the practice of English statisticians like Galton and Karl Pearson, Spearman now proposed to apply the exact method of correlation to psychological data, much in the same way as it had already been applied to the physical data of biometry and anthropometry; and, devising a short simple and intelligible "footrule" for calculating such coefficients, he did much to popularise the use of correlation among non-mathematical psychologists[3].

---

[1] *Psych. Arbeiten.* II., p. 324.

[2] It should be added that for the measurement of simple and specific capacities, such as sensory discrimination, reaction-time, and the like, tests involving elaborate apparatus still remain indispensable. These more elementary capacities, however, are but remotely related to the work of the classroom; and most of the testing of this type is still carried out, not in schools upon children, but in psychological laboratories upon adult students. On the use and technique of such tests an excellent handbook is provided by Whipple's *Manual of Mental and Physical Tests.* Standard methods are there described in full. Detailed results are brought together; and complete references to the literature are appended for each test.

[3] *Amer. J. Psychol.*, XV, 1904, pp. 72 *sqq.*; *Brit. J. Psychol.*, II, 1906, pp. 89 *sqq.* A coefficient of correlation is a fraction or percentage expressing the amount of agreement between two series of measurements: for detailed explanations and examples, see Appendix V.

Previous investigators had trusted mainly to the coarse method of mere inspection. They made rough classifications of their examinees into bright and dull; glanced at their experimental data; and then declared that, according to their general impression, there was, or was not, a correspondence between one test and another, or between the tests, on the one hand, and the teachers' estimates of ability, on the other, and that the correspondence was high or low. A procedure so subjective, though to this day too frequently relied upon by teachers and educationists, can hardly claim to be scientific. Little wonder that the earlier inferences were so divergent and conflicting. But this loose guessing was now to be abandoned. The use of a sound statistical method of comparison made it possible to ascertain conclusively the existence of a correlation between two varying qualities, apart from all personal bias, and to measure such correlations objectively, so that they themselves could be validly compared. When previous data were re-examined in the light of these more accurate devices, and the new methods applied to the old results, it could often be shown that the original deductions of various experimenters, not only contradicted each other, but were actually opposed to the true significance of their own figures. In this way, many of the bewildering discrepancies could immediately be resolved. By the same means, too, irrelevant sources of error could be gauged and eliminated: and, after this had been achieved, an appreciable correlation would frequently emerge, where none before had been suspected.

15. *The Conception of General Ability.* In Professor Spearman's hands the adoption of this new statistical weapon led at once to a remarkable triumph. Although at first attacked with considerable vigour, his generalisations have opened up an entirely fresh field and have initiated an entirely new campaign. Of his various conclusions the most important is the theory of a general ability underlying all the various mental activities that admit of being tested.

Dr. Johnson once declared that, had Newton applied himself to poetry, he would have written a great epic. A critic immediately objected that one man might have great learning, another keen judgment, another a fine imagination. "No, sir," replied Dr. Johnson, "it is only that one man has *more mind* than another. He may direct it differently; he may by accident desire to excel in this study or in that. Sir, the man who has

vigour may walk to the east, just as well as to the west."[1] The same belief in the versatility of genius, in the diffusion of defect when defect is present, underlies Professor Spearman's hypothesis of a central fund of mental energy. His formulation of the hypothesis, however, was no mere ingenious conjecture. It was the outcome of an elaborate statistical analysis, applied to a long series of experimental tests.

In England several researches were immediately commenced to test his methods and results. His mathematical procedure and his psychological conclusions have alike been sharply debated[2]; but, on the whole, the final trend of recent research has been more and more to corroborate his leading principles. At Oxford and at Liverpool, for example, early investigations[3] with tests of increasing complexity, intended to tap both lower and higher levels of the mind, appeared to verify the hypothesis of a so-called central factor, of a general ability, pervading not only simple sensory functions, but also radiating in different directions and to different degrees among all intellectual capacities. At the same time it was shown that the capacities reached by the more searching tests were apparently for the most part hereditary or congenital; and that, unlike the ordinary scholastic examination, these devices were measuring not so much acquired capacity—knowledge gathered by memory, or dexterity gained

[1] Boswell, *Journal of a Tour to the Hebrides*, Aug. 15 (Carruthers' ed., p. 16). The distinction between general and special intelligence was recognised by Aristotle (*εἶναι δέ τινας δοφοὺς οἰόμεδα ὅλως, οὐ κατὰ μέρος. Nic. Eth.* VI. vii. 2).

[2] The technical criticisms of Spearman's doctrines cannot be discussed in detail here. Dr. William Brown and, more recently, Professor Godfrey Thomson (*The Essentials of Mental Measurement*, Revised Edition, 1921) have criticised the validity of the mathematical formulæ used; Professor E. L. Thorndike (*Amer. J. Psychol.*, XX., p. 364) and his pupil Simpson (*Correlations of Mental Abilities*) have criticised the doctrine of a general factor. Dr. J. C. Maxwell Garnett has intervened mainly to the support of Spearman (*Brit. J. Psychol.*, 1919, IX., iii. and iv., 1920, X., ii. and iii.). From later contributions upon either side, it now appears that the differences of view are by no means so complete or irreconcilable as they once appeared. Most of the misgivings seem mainly to concern the adequacy of the theoretical proofs offered. When it comes to practice, the most recent and the most cautious of the critics is found adopting the same working hypothesis and employing much the same tests for general educational ability (*e.g.*, Thomson, *Brit. J. Psychol.*, 1921, XII., iii.).

[3] Burt, *Brit. J. Psychol.*, 1909, III., i., *J. Exp. Ped.*, 1911, I., ii.

through practice—but rather something inborn. The two conclusions seemed to suggest a serviceable working definition for the vague term "intelligence." It was accordingly defined as *inborn, general, intellectual efficiency.*

VIII.—TESTS OF HIGHER MENTAL PROCESSES.

16. The results of the researches just described showed that, for purposes of practical testing, complex processes were more important than simple. They agreed, indeed, with the earlier declarations that simple sensory and motor capacities were dependent upon intelligence; but at the same time they proved quite clearly that the degree of this dependence was comparatively small, so small in fact that it might easily be obscured by errors of measurement or crudities of analysis. With complex abilities it was different. It was now plain that from them far better results could be obtained. Higher mental processes, those, for instance, in which both sense perception and motor activity were combined, furnished tests far more effective than the lower; and, as a general rule, it was discovered that *the higher and the more complex the activity tested, the closer was the correlation with intelligence.*

Instead, therefore, of measuring intelligence by the speed of simple tapping, the investigators now required the examinee to tap an irregular row of dots (the task involved in using McDougall's "dotting machine")—each tap demanding a distinct effort of aiming and the whole experiment calling forth a high degree of sustained voluntary attention. Instead of asking him to draw a second line equal to the first, or to bisect a single long line, he was required to divide it into two parts in the same proportion as the two parts of a smaller line already divided. Instead of asking him merely to say any word associated with the given test-word, he was required to name some word standing in a specified logical relation to the test-word—an opposite, a synonym, a whole of which it formed a part, or a genus of which it formed a species; for example, to "Black" he must reply "*White*"; to "Bad," "*Evil*"; to "Leg," "*Body*"; or to "Dog," "*Animal.*" Sometimes the logical relations were not single and uniform, but manifold and mixed, as in the well-known test of *Analogies.* In this test, the examinee was shown three words, the first pair indicating a definite logical relation; and was required to work a sort of

" rule of three " in words instead of numbers ; for example,

1. Black *is to* White *as* Bad *is to* . . . ?
2. Leg *is to* Body *as* Wheel *is to* . . . ?

Sometimes the logical relations had to be combined to form inferences, as in the so-called test of *Syllogisms* and in the many other tests of reasoning ; the following are simple illustrations :

1. Tom is taller than Jim ;
   Jack is taller than Tom.
   Which is the tallest of the three ?
2. John said : " If the train is late, I shall miss my appointment ; if the train is not late, I shall miss the train."
   We do not know whether the train was late or not.
   Can we tell whether he kept his appointment ?

The foregoing instances may be regarded as simple tests of constructive reasoning—the earliest and the most commonly used type. Tests of critical or destructive reasoning were almost equally effective ; and have recently formed the basis of an ingenious scale by Dr. P. B. Ballard[1]. Tests of this latter type are usually termed Absurdity Tests. Binet had already introduced such a test, consisting of five absurdities to be explained by children aged eleven. The child was asked :

1. " ' I have three brothers, Jack, Tom, and myself.' What is silly and absurd in that ? "
2. " ' Once the body of a poor girl was found in a wood, cut into eighteen pieces. They say that she killed herself.' What is foolish in that ? "

Binet's three remaining " absurdities " contained similar inconsistencies, of a rather gruesome kind to set before young children.

But five questions do not make a scale. Accordingly, for this as for so many other tests, in order to multiply the amount of possible marks, investigators have adopted one of two possible devices ; they have either compiled a number of discrete questions or statements, or else combined the whole into one consecutive test passage. Dr. Ballard, in his Absurdity Test, employed the former method ; and collected as many as thirty-four nonsensical statements, which he graded in a series of ascending difficulty. Others preferred the second method, illustrated by the

[1] " The Limit of the Growth of Intelligence," *Brit. Journ. Psychol.* XII., ii., (1921), pp. 129–131.

extract which follows. The child is required to read the passage and to discover as many absurdities as he can :

> *"A Sunday in France.*
>
> "Ten years ago on a pleasant summer's afternoon in the middle of January, 1916, the twelve o'clock express from Scotland was rushing past the busy terminus of the Great Western Railway at twelve miles an hour.
>
> A clean-shaven young Englishman, of about fifty years of age, stepped lightly from one of the first-class carriages and hurried slowly down the platform, with both hands in his pockets, carrying a heavy bag, and gaily curling the tips of his moustache.
>
> There was not a cloud in the sky; and as the rain was still falling heavily, he took off his mackintosh and strolled out into the crowded streets of the city. The ripening fields of corn through which he passed were turning golden as the sun set in the south :" etc.

Tests of these various types were found to correlate closely with intelligence. It will be observed that they depend upon no special knowledge ; the words that have to be read by the examinee are well within the powers of children at the ages specified. Hence, the answers to the questions depend, not upon ability to read, but upon ability to think, not upon acquired information or skill, but upon native capacity of intellect[1].

The better tests, therefore, seemed, as a result of all the different experiments, to be those requiring for their performance, *the higher mental processes* : and the best tests of all, those involving the highest levels of thinking—*the ability to reason.* Henceforward, for the testing of general intelligence simple sensory or motor capacities have been all but universally abandoned. Except, perhaps, when dealing with the youngest or dullest children, no one would now attempt to gauge educational capacity by simple measurements of skin-discrimination or speed of tapping.

---

[1] Of researches upon reasoning processes those of Mr. W. H. Winch in England, (*Brit. J. of Psych.*, 1914, VII. p. 190, and *Journ. of Exp. Ped.*, VI. 1921, p. 121), and of Professor F. G. Bonser in America (*Columbia Contributions to Education*, 1910, XXXVII), merit notice. A scale of graded reasoning tests for English school children was published in *Journ. Exp. Ped.* V. ii and iii (1919) pp. 68 *et. seq.* ; and an abridgement is reprinted in Appendix VIII.

17. *The Possibility of Group Testing.* At the same time, these newer investigations showed, what had formerly been denied, that group tests might prove just as trustworthy for the measurement of intelligence as the individual tests which up to now had principally been favoured[1]. Tests of sensory or motor capacity, indeed, are difficult to apply to a number of individuals simultaneously; but tests of these higher and more complex capacities can readily be carried out as written class exercises: and the results so procured are both self-consistent and highly diagnostic. Of this conclusion one immediate result was that theoretical inquiries into the relations between various intellectual processes could be carried out much more rapidly and upon a most extensive scale. The practical corollary, the value of group-testing for everyday purposes, was not developed until a later period in the history of mental tests.

18. *Specific Abilities.* Subsequent researches by Spearman and his followers led to the distinction of special capacities side by side with general ability; and the most recent investigations seem to show that any one concrete intellectual activity—such as memorising a given poem or working a given set of sums —may be considered to depend upon intellectual factors of three different orders: first, the *general factor*, common to all intellectual activities, and known usually as general intelligence; secondly, one or more *special or "group" factors*, shared only by a limited number of intellectual processes; and, thirdly, *specific or individual factors*, peculiar to each particular test itself. The difference, no doubt, is principally a difference of degree. It seems possible to regard the "general factor" as simply the "group-factor" that is of the most widespread occurrence: and the "specific factors" as simply the "group-factors" that are most narrowly limited in their operation.

In the ordinary work of the Elementary School the more important "special" capacities (if one may trust the results of experiments on a somewhat limited scale) were shown to be the following: (1) *arithmetical* ability; (2) *linguistic* ability (which perhaps may be separated into two types—(*a*) the more elementary *verbal* factor entering into such simple activities as reading and spelling, and (*b*) the more highly developed *literary* factor entering into the activities of English composition); (3) *manual* ability (which no doubt has also several subordinate

[1] See *J. Exp. Ped.*, 1911, *loc. cit. sup.*

forms); and perhaps also (4) *artistic* ability and (5) *musical* ability[1].

Special abilities of a more elementary and strictly psychological kind have been found far harder to determine. As already observed, in supposing that definite faculties could be measured by one or two typical tests, the earlier experimenters proved to be much mistaken. The few recent inquiries that have approached the problem with the proper statistical methods—those of correlation and "partial" correlation—have so far been unfruitful: they have as yet succeeded in isolating no special unitary functions; much less have they succeeded in devising for such functions any suitable specific tests[2].

19. *Mental Imagery.* One early attempt, however, in the analysis of special abilities—qualitative rather than quantitative in its method—has of late fallen into somewhat undeserved neglect. Galton had classified men according to their predominating type of mental imagery, distinguishing the "visualiser" from the "audile," and both from the "motile," and those again who think in visible pictures of concrete things from those who think chiefly by means of heard or uttered words.

---

[1] See *Distribution of Educational Abilities*, pp. 56–63. Reference should here be made to the early work of Mr. R. C. Moore in Dr. Burt's laboratory at Liverpool, of Mr. Bradford in Dr. William Brown's laboratory at King's College, London, and particularly of Miss Nellie Carey in Professor Spearman's laboratory at University College, London; a full account of her researches in London County Council Schools will be found in her articles on "Factors in the Mental Processes of School Children," especially Part II. "On the Nature of Specific Mental Factors," and Part III. "Factors concerned in School Subjects." (*Brit. Journ. Psychol.* 1916. VIII., i and ii, pp. 70–92, 170–182).

[2] A striking instance is the negative results of experiments attempting to demonstrate the existence of a special motor ability—one which casual experience would suggest was among the most easily proved. After a careful series of tests, carefully compared and analysed, Professor Muscio concludes: "There can be no general motor test because there is no 'motor type.' Motor capacities appear to vary independently of one another" (*Brit. Journ. Psychol.* 1922. XIII. ii, p. 184). It should be added, however, that the most recent investigations carried out in the psychological laboratory at Cambridge, and at present unpublished, indicate that this apparent independence of motor capacities only holds good on the simplest and lowest levels—a conclusion quite in accordance with the earlier researches on simple motor tests alluded to above, in § 9.

For this purpose he used a test based on the familiar device of a standardised questionnaire. He asked his subjects to think first of some definite object, for example, " your breakfast-table as you sat down to it this morning," and to consider carefully the picture that rises before the mind's eye. " Is the image dim or fairly clear ? Are the colours of the china, the toast, the mustard, meat, or parsley, or whatever was on the table, quite distinct and natural ? " " Can you see mentally more than three faces of a die, or more than one hemisphere of a globe at the same instant of time ? "

" Can you recall with distinctness the features of near relatives and other persons ? " " Can you hear in your mind's ear a note which is too high for you to sing ? " " Can you, with your lips open and your teeth apart, think mentally of such words as ' bubble ' or ' putty ' ? " " Can you in imagination hear the clinking of teaspoons, the slam of a door ; or smell the odour of tar, of an oil lamp just blown out ; or taste sugar, chocolate, lemon juice, or currant jelly ? " According to the nature of their replies, Galton graded different persons, for each form of mental imagery, into what he termed eight octiles, ranging from those whose imagery was " brilliant and distinct," to those whose power of concrete imagination was " practically zero."

The detection of such " imaginal " types was at one time thought to be of great significance for education. The visualising child was to be educated upon " look-and-say " principles ; the audile by " phonic " methods. But the first enthusiasts, by their wild deductions and excessive claims, brought all such efforts into discredit.

With other mental functions there have been similar endeavours to classify individuals according to qualitative types —for example, in memory and in attention ; but these have stimulated less interest, and have led to little experimental work[1].

There can be small doubt that, if the teacher were able not only to measure his pupils' intelligence, but also to discover the different qualitative peculiarities of their minds, the individualisation of teaching methods would be enormously enhanced.

---

[1] See for a summary of the chief contributions Meumann, *Experimentelle Pädagogik*, Xte Vorles., Stern, *Differentielle Psychologie*, Kap. VI., XIV. A convenient English summary will be found in Rusk's *Experimental Education*, Chap. XI.

The practical value of such analysis has lately been demonstrated in the case of children with special talents or special disabilities. As recent surveys both in Birmingham[1] and London[2] seem to have shown, the backward child is often backward only because of some special disability. The so-called "word-blind" child, for example, proves often to have been handicapped, under a "phonic" method of teaching, by a poverty of auditory imagery, or, under a "look-and-say" system, by a poverty of visual imagery. Similarly defects in the various forms of memory will produce grave backwardness in such fundamental subjects as spelling and arithmetic. When these underlying defects have been diagnosed, and the method of instruction appropriately changed, the backwardness (as the published records of individual cases show) may vanish in a year[3].

On the whole, however, it must be admitted that the measurement of the special elementary functions of the mind as yet has hardly begun. In their diagnosis the psychologist has to rely upon observation rather than upon testing. General intelligence has been found in comparison so easy to test, and of such widespread significance, that the testing of special functions has been largely passed over.

IX.—INDIVIDUAL TESTS: THE BINET-SIMON SCALE.

20. Meanwhile abroad, a long succession of inquiries, stimulated largely by the work of Dr. Alfred Binet, in France, had been tending, though by a wholly different route, to conclusions closely resembling those reached in England by the use of correlation. As early as 1895 Binet had published a programme for studying the relations existing between different psychological processes[4]. By means of ten tests for ten separate functions

[1] B. R. Lloyd and C. Burt. *Report of an Investigation upon Backward Children in Birmingham.* (City of Birmingham Stationery Department, 1921).

[2] C. Burt, *Mental and Scholastic Tests*, L.C.C. Reports, 1921.

[3] See Bronner, *Psychology of Special Abilities and Disabilities*, S. L. Hollingworth, *Special Talents and Defects* and Lucy Fildes, "Word-Blindness," *Brain*, XLIV. iii., pp. 286–307.

[4] *L'Année Psych.* (1896), p. 411.

he hoped to secure, in the space of one or two brief hours, an exhaustive survey of any given personality. The practical success of Bertillon in determining so-called "physical constants" by a few rapid measurements of individual criminals, seems to have aroused high expectations of achieving a "psychic portrait" by methods no less simple and summary.

21. *The Diagnosis of Mental Deficiency.* Binet's Scale was first constructed for examining the intelligence of Parisian school children suspected of mental deficiency and recommended for transfer from the ordinary school to special classes; and it is for this purpose that the Scale has been increasingly used. At the suggestion of the late Professor J. A. Green, Miss K. L. Johnston went over to Paris to study the new methods; and on her return contributed an English translation of Binet's 1908 Scale to Professor Green's journal[1]. In the same year, Dr. F. C. Shrubsall published a critical account of the tests, and of their uses in the examination of children for special schools[2]. In the following year, the Annual Report of the Chief Medical Officer for the Board of Education contained a summary of the entire Scale; and the tests were expressly recommended for use in cases of suspected mental defect[3].

22. *Simplified Methods.* Binet began, as we have seen, with simpler tests of motor and sensory capacity, such as tapping, skin discrimination, and the measurement of reaction-times. But, at a very early stage, he felt the need for testing processes of a more complex character, such as might approximate more nearly

---

[1] *Journ. Exp. Ped.* I., i. (1911) pp. 25 *et seq.*

[2] *School Hygiene* (1911), pp. 613 *et seq.*, "The Examination of Mentally Defective Children." Reference should also be made to the pioneer work carried out by the school medical officers in this country, particularly by Dr. James Kerr (formerly School Medical Officer and now Medical Research Officer under the London County Council) and his various colleagues. Their investigations, may be found incorporated from time to time in past Annual Reports of the Medical Officer (Education) published by the London County Council. The more recent studies by Dr. E. O. Lewis, published in the *Journ. Exp. Ped.*, IV. iv. (1918) pp. 198–202, and in *Studies in Mental Inefficiency* IV. ii. (1923) pp. 27–34 should also be referred to in the same connection.

[3] *Annual Report for* 1912. Appendix E. "Schedule of Medical Examination of Children for Mental Defect."

to the practical activities of ordinary life. He proposed, accordingly, to reject what he described as "the brass instruments of the band of German psychologists"; to exchange the laboratory for the school; and to use for his experiments "no apparatus except pen, paper, and a little ink."

In 1896 he began by examining 80 children, asking them to describe a simple picture, and classifying the descriptions into four or five generic types. He claimed that his conclusions formed "the first achievement hitherto obtained from an experimental study of the higher intellectual faculties[1]." These early experiments gradually led him to formulate a definite view of intelligence, which, though not published in full until a later date, guided the course of his main researches.

He began by distinguishing natural intelligence and acquired culture. And then, contrasting the behaviour of the normal intelligent child, on the one hand, with that of unintelligent defectives, and of the demented and insane, on the other, he was led to separate mental processes into two kinds or levels—the lower and the higher. The lower processes consist of the simple reception of sensory impressions and (as he phrased it) the "mechanical unchaining" of habitual associations. The higher processes are characterised by attentive co-ordination, manifested in three ways—through purposive direction, through active adaptation, and through conscious correction. The former are analytic; the latter synthetic. The former reproductive; the latter creative. The former are automatic and unregulated; the latter are controlled and self-critical, and exhibit progressive learning through trial and experience. These, in Binet's view, are the chief marks of intelligence.

In any intelligent process, he adds, "the greater number of our elementary faculties are involved. . . . Nearly all the phenomena with which psychology concerns itself are phenomena of intelligence." Hence, a test of almost any faculty is, in a sense, a test of intelligence; and, to yield a satisfactory measure, not one test, but many different tests in combination, must be used. "To cover a wide field of observation, it goes without saying that our tests must be manifold and heterogeneous"; and, "instead of measuring the intensity of simple faculties, the vain ambition of the psycho-physicists, we shall measure acts of adaptation. Understanding the normal progress of development,

[1] *L'Année Psych.*, (1897), p. 296.

we shall be able to determine how many years an individual is advanced or retarded[1]."

Though sometimes known as the "French" procedure,[2] in contrast to the "German" (which still clung, for the most part, to the traditional class of laboratory experiment), the new type of experiment was not exclusively confined to France. Two years after Binet's first contributions, Ebbinghaus, influenced in all probability, by Binet's fresh method of approach, attacked the problem from a standpoint entirely new in Germany. Being requested by the educational authorities of Breslau to assist in an official examination of certain schools of Silesia, he proceeded to invent a new test of intelligence, which he called a "Kombinations-Methode" (completion-method)[3].

Like Binet, he began by analysing the real psychological nature of processes which men in ordinary life are accustomed to regard as intelligent: such processes, he thought, consist invariably in "bringing together a multitude of separate and disordered fragments of experience into a single whole that is unitary, full of meaning, and, in some way or other, purposive." Like Binet, therefore, he concluded that intelligence is exercised more in synthesis than in analysis, more in the re-organisation of sensory percepts than in their bare discrimination. Intelligence, in a word, is essentially a process of "combination[4]."

It was this synthetic process, accordingly, which he aimed at eliciting by his new test. A mutilated passage of prose was

---

[1] See *The Development of Intelligence in Children*, pp. 40 *et seq.*, 253 *et seq.*, and *The Intelligence of the Feebleminded*, Chap. XIII., "A Scheme of Thought" p. 130 *et seq.*

[2] See Sharp and Titchener (*Amer. Journ. Psych.*, X., pp. 348 *et seq.*), who expressly set out to investigate this new departure, but with no very promising results.

[3] *Zeitschr. f. Psych. und. Phys. d. Sinnesorg.*, XIII. (1897), p. 401. This important paper of Ebbinghaus has been described as "the first great attempt to grapple with the problem." But, as Binet himself points out, experimentation on these new lines in Germany, notably the Würzburg School's researches upon the psychology of thought, were really stimulated by his own early researches upon intelligence. And, though Binet's 'scheme of thought' was not published in explicit form until 1909, its origin and conception "dates" (as he says) "much further back."

[4] It is impossible not to relate this view of mental activity with Sir Charles Sherrington's conception of the "*Integrative Action of the Nervous System*" as a whole. See his Yale lectures published under that title (1906) and earlier papers there cited.

put before the examinee, and the latter was required to fill in the missing words, thus combining the disconnected words into an intelligible whole, by completing the sense. Between the order furnished by his test, and the order of ability in school work, he found a correspondence remarkably close.

Ebbinghaus' procedure, though for long neglected, has since proved of much importance in the history of mental testing. His technique is adaptable to tests of very different types; and of late has even been applied for tests of special knowledge and information. The following specimen is from the opening paragraph of a completion-test widely used in this country:

> "Revenge is a . . . of wild justice, which the more man's nature runs to, the . . . ought law to weed it out. For, as for the first wrong, it . . . but offend the law . . . the revenge of that . . . putteth the law out of office[1]."

The earlier completion-tests were based upon simple narratives or stories, depending largely upon imagination and visualisation for their understanding. A test, which, like the foregoing example, depends upon the power to follow and complete a logical argument, has been shown to yield, at any rate with older children, a far better measure of intelligence[2]; and it has even been ascertained that, for young children, the principle of the completion-or combination-method can also be adapted to pictorial instead of verbal material, with almost equally satisfactory results[3]: the test of "reconstructing dissected pictures" (as it was then termed) has more recently been developed in various forms[4] and freely used as a test of the so-called "performance" (or non-linguistic) type.

In America, as has just been observed[5], the first attempts to use these simplified methods, in accordance with the "French"

[1] From the first sentences of Bacon's *Essay on Revenge*.

[2] The Consultative Committee is, of course, of opinion that the test is of greater value if it does not offer any scope for memory. They recommend that sentences should be specially framed for this purpose. The Committee also suggest that tests involving reasoning power are more useful than those of mere pictorial representation.

[3] *Journ. Exp. Ped.*, I., ii (1911), p. 102.

[4] For an excellent instance of this test, derived partly from the principle of the old 'form-board' test, see Healy, *Psych. Rev.* (1914), "A Pictorial Completion-Test"; and for a critical survey of work done with tests of this character see Pintner and Anderson, "The Picture Completion Test" (*Educ. Psych. Mon.*, Warwick & York).

[5] See footnote [2], p. 25.

procedure, were not encouraging. But later researches, both in England and in the United States, produced a large series of simple tests for higher mental processes, which played afterwards a considerable part in the examinations by the written or "group" method[1]. Tests of this character, however, were for a while cast into the shadow by the great interest, aroused in America and elsewhere, by what is now known as the Binet-Simon Scale.

23. *Intelligence Measured by an Age-Scale.* Binet's most celebrated achievement was the construction, upon a plan entirely novel, of a metric scale (*échelle métrique*) for the measurement of intelligence. The final conception of this scale he reached only by steps and stages, so that there is in reality, not one Binet scale, but a series of several scales. In its latest revision the fundamental purpose of the scale was to measure the development of intelligence in terms of mental age[2].

With his new unit of measurement, his new and simplified methods, and his new and suggestive definition of intelligence, Binet now abandoned his earlier attempts to extract a quantitative grading in marks from a single consecutive test; and assembled instead a miscellany of test-problems of the most varied character. A correct answer to each test-problem was to count as an equivalent fraction of a mental year. His starting-point, as we have seen, was not the usual experiments of the laboratory psychologist, but rather the empirical method of the teacher and the alienist. Teachers and doctors had long been accustomed to examine their scholars or their patients by means of questions of a simple conversational type, questions meant for the most part to show the state of mind through general information or scholastic knowledge. This form of oral interview Binet adapted and improved by standardising

[1] Among earlier American researches with tests of this type, the following deserve mention: Sharp and Titchener, *Amer. J. of Psych*; p. 348. Aikins, Thorndike and Hubbell, *Psych. Rev.*, IX., 1902, pp. 374 *et seq.*—both mainly negative in their results. Of those yielding more positive results, Whitley, "Tests for Individual Differences," *Arch. of Psych.*, 1911, XIX., is one of the more important. A model attempt at standardising association tests of this kind is that of Woodworth and Wells, *Psych. Mon.*, XIII., 1912.

[2] The concept of a mental age, though new to the psychologist, was not (as we shall find) new to the educationist. See below, § 34, on the standards formulated by the codes of the Education Department at Whitehall in the sixties of last century.

and grading the questions to be put. He asks the child, for example, such simple questions as these: "What is your name?" "Are you a little boy or a little girl?" "How old are you?" If the child answers the first two correctly, but fails to answer the last, the examiner can already infer that the child's mental age is probably about four years, since a normal child of three can usually give his name and sex, and a child of five can state his own age. Of the various questions that make up the scale, many, like the three just cited, and the recognition and naming of four primary colours (a test for age five), of common articles, such as a knife, key and penny, and of objects in a picture (tests for age three), and of the four commonest coins (a test for age six) pre-suppose merely the normal amount of parental instruction that would be given in an ordinary home. Others depend chiefly upon the teaching of the ordinary school, as naming the days of the week, the months of the year, and the date, or the exercises in reading, writing, dictation and simple computation. Others, again, such as the repetition of numbers progressively increasing in length—32, 714, 8596, and so on—hardly depend at all upon experience or instruction. Finally, many of the tests still recall the old laboratory experiments upon sense-discrimination: for example, the comparison of graded weights, and of lines differing in length.

Binet's last version of his scale (1911) contained fifty-four such simple tasks or questions, five being allocated to each age of school life from three to the adult stage, with the exception of years XI., XIII. and XIV. The child was assigned a mental age corresponding to the hardest group of questions correctly and completely answered. For the correct answer to any additional question suitable to a higher age, one-fifth of a year was added.

Binet, and his collaborator, Simon, seem originally to have chosen their tests to gain information, not only about general intelligence, but also about special functions: they speak of such "faculties" as sensorial perception, attention, memory, comparison, abstraction, and so forth. But later, this possibility seems largely to have retreated into the background; and certainly such light as the Scale yields in this direction is but a general dim glimmer with an occasional illuminating flash. Probably the most striking and the most unimpeachable principle involved in the Scale is that of the heterogeneity of the test-problems. If, as we have found, every test-process depends, not only upon general intelligence, but also upon specific

capacities—or "faculties" (as Binet calls them), it follows that intelligence can never be measured accurately, unless the tests are sufficiently numerous and diverse to eliminate, by averaging, the varying influence of these specialised factors. Since Binet's day it has been a cardinal principle of intelligence-testing *never to employ* (as had hitherto been done) *any one test singly*, but to combine tests of different types into an average measure.

Binet did not check the validity of his scale as a whole, nor yet of its component tests, by the method of correlation, but this method was soon afterwards applied; and it was found that, at any rate with younger and subnormal children, the total results correspond very closely with independent estimates of intelligence. Not until quite recently, however, was the diagnostic value of the separate test-problems investigated by an adequate statistical method. It was then found that the several problems differed enormously in their worth; some, such as the test of so-called suggestibility[1], being of no significance whatever.

24. *Revisions of the Binet-Simon Scale.* Binet himself, as we have noted, published many different versions of his own Scale[2].

---

[1] The child is shown three pairs of lines increasing in length, the right being always longer than the left; he is then shown three pairs of lines of equal length. The dull suggestible child—the child who acts automatically upon any guiding idea that is suggested to him—still continues (or is expected to continue) mechanically to declare the line on the right is longer, even when no difference is in fact perceptible.

[2] The following are the chief articles published by Binet upon this subject:—

(1) Upon the Necessity of Establishing a Scientific Diagnosis of Inferior States of Intelligence. (*L'Année Psych.*, 1905, pp. 163–191).

(2) New Methods for the Diagnosis of the Intellectual Level of Subnormals. (*L'Année Psych.*, 1905, pp. 191–244).

(3) Application of the New Methods to the Diagnosis of the Intellectual Level among Normal and Sub-normal Children in Institutions and in Primary Schools. (*L'Année Psych.*, 1905, pp. 245–336).

(4) The Development of Intelligence in the Child. (*L'Année Psych.*, 1908, pp. 1–90).

(5) New Investigations upon the Measure of the Intellectual Level among School Children. (*L'Année Psych.*, 1911. pp. 145–201).

Other small changes appeared when the tests were re-described in *Les Idées Modernes sur les Enfants* (1908) and in the *Bulletin* published through *La Société Libre pour L'Etude Psychologique de l'Enfant.*

In the first form (1905) the tests were grouped according to stages rather than ages. It was only in the 1908 version that Binet showed the practicability of measuring intelligence with a single year as a unit. In the 1908 version, however, the number of tests assigned to successive ages differed greatly: and in the 1911 revision he decided to assign five standardised tests to every year. Even this revision, usually accepted as Binet's final version, was printed and reprinted in different books and pamphlets with minor modifications; and, almost to the day of his death, he was still engaged upon its revisal. It is clear that he did not in any way regard his published versions as ultimate.

25. (i) *The Vineland Revision.* In America the Binet Scale was welcomed with much eagerness. Dr. H. H. Goddard (at that time Superintendent of the Training School for Defectives at Vineland, New Jersey) issued an early translation of the Scale[1]; and introduced a few slight changes, chiefly by adding one or two fresh tests (the well-known code[2] and clock[3] tests —the latter borrowed from an earlier publication of Binet).

This translation was for long the standard version in use in the United States; and has formed the basis of nearly every American adaptation since. But the original claims of virtual infallibility, at first made by American enthusiasts, were not borne out by subsequent experience; and the need for a far more thorough revision, particularly in respect of the age-assignments of particular tests, was more and more acutely felt.

Of the later versions compiled by American psychologists, two are so drastic as virtually to constitute new scales, and deserve a brief description here.

26. (ii) *The Yerkes Point-Scale.* The first is the Point-Scale of Professor R. M. Yerkes. This consisted of twenty exercises picked, with one exception, from the original Binet-Simon series.

[1] *The Binet-Simon Measuring Scale : Revised Edition.* Vineland, 1911.

[2] This test consists in translating from memory the words " come quickly " into the proper symbols, letter by letter, according to a simple diagram code, said to have been used during the American Civil War.

[3] This test consists in interchanging in imagination the positions of the large and small hands of a clock—supposed to point originally *e.g.*, to twenty-two minutes past six—and stating the time then indicated.

The addition was a test first used for the measurement of intelligence in England, and then named the Analogies Test. Many of the less satisfactory tests were discarded by Yerkes; and the results were computed not in terms of mental years, but in terms of simple points or marks. Instead of marking the results by the all-or-none, pass-or-fail method of Binet, partial credit was also given for various attempts according to their different merit. From a theoretical standpoint, the work done by Yerkes and his collaborators was richly suggestive. But for the two chief practical purposes of the Scale—the examining of borderline defectives at the age of admission to special schools, and the examining of supernormal children at the age of sitting for Junior County Scholarships—the Point-Scale appears to be, at any rate for English children, in no way superior to the original[1].

27. (iii) *The Stanford Revision.* The second and more recent version is the Stanford Revision and Extension, drawn up by Professor L. M. Terman and his collaborators in California. The Stanford Scale contains ninety tests. For most of the years there are six tests instead of five, so that intelligence is measured in terms of months, instead of in terms of decimals of a year. Of this revision the salient merit lies in the addition of many excellent tests, both new and old, principally for the higher ages. Terman, for example, asks the child to tie a bow-knot, to interpret fables, to explain with a pencil on a little plan how he would search for a lost ball in a circular field, to explain resemblances between things as well as their differences, to repeat numbers backwards as well as to repeat them forwards. One of his best-known tests is to define the meanings of a graded vocabulary of 100 words, a test applicable at almost any age, and, according to Terman, the most effective in the whole list.

Some of the new tests, however, still have a scholastic bias. For example, a test for age 14 consists of three arithmetical problems, such as the following:—

> "If a man's salary is 20 dollars a week and he spends 14 dollars a week, how long will it take him to save 300 dollars?"

---

[1] For an excellent discussion of the Point-Scale, with results obtained from English low-grade children, see Dr. E. O. Lewis, "The Binet and Point-Scale Methods of Testing Intelligence" *Journ. Exp. Ped.* iv. (1918) pp. 198–202.

And again, among the tests for superior adults there is this problem :—

> "A mother sent her boy to the river, and told him to bring back exactly 7 pints of water. She gave him a 3-pint vessel and a 5-pint vessel. Show me how the boy can measure out exactly 7 pints of water, using nothing but these two vessels and not guessing at the amount. You should begin by filling the 5-pint vessel first."

Unfortunately, too, Terman still retains many of the original tests of Binet and Simon, that have since been proved to be of little value. Hence the application of the Stanford Revision consumes even a longer amount of time than the original version. During the last year or two the Stanford version has been widely employed in this country, but no re-standardisation has yet been published, specially adapted for English children. Careful experiments show that the age-assignments given by Terman, though more suitable than those of Binet, are still in many cases inexact for children attending English schools ; and, with the sanction of Professor Terman himself, a further re-standardisation of these tests is now being carried out in this country[1].

28. (iv) *The London Revision.* Meanwhile, an English re-standardisation of the original Binet-Simon Scale has been already published, based upon extensive experiments carried out in London schools. It was felt that, for the present at any rate, two alternatives were possible ; first, to take the original version of Binet and Simon, which has formed the foundation of all subsequent work, and which has been used so extensively by school medical officers in this country, and re-standardise it as it stood ; secondly, to carry out a wholesale reconstruction of the scale according to some entirely new scheme. Merely to revise the original selection by adding three or four new tests, and discarding one or two of the old, seemed of little service. A radical revision, however, must be the work of years. Hence, for provisional purposes, in compiling what may be called the London Revision[2] it was decided to adhere as closely as possible

[1] For further observations on this Revision, see the Committee's comments below, §59.

[2] The Committee points out that this adaptation of the Binet-Simon scale for English children was arranged by Dr. Cyril Burt in consultation with Dr. Th. Simon. It is set out, with a useful commentary, in Burt's *Mental and Scholastic Tests*, London County Council Reports, 1921.

to the actual methods of Binet and Simon themselves. The only considerable alterations that have been made consist in a re-assignment of the individual tests to the more appropriate ages, as required by the results of careful experiments upon London children. The general frame-work of the original scheme, and all the original tests, have been retained : so that the results can be computed either in a form directly comparable with the original Binet age-arrangement or in a form providing measurements, as accurate as possible with such a scale, for English children in English schools.

29. (v) *The Treves-Saffiotti Method.* A modification of the Binet-Simon Scale, which, among English-speaking investigators, has attracted but little recognition, is that elaborated in Italy by Professor U. Saffiotti and Professor Z. Treves. The changes proposed are principally two : first, the tests are to be grouped not only by age, but also by school-class : secondly, the children tested are to be graded by description of mental quality instead of by arithmetical computation of years or marks attained. For children of a given age in a given class there are allotted three sets of tests—easy, medium, and hard. According to their success in these, the children are grouped as *deboli*, *medii*, and *forti*—dull, average, and able —and marked D, M, or F. If time allows, the children may be tested with other sets of tests besides those immediately suited to their level ; and, according to their further success in these, the children in each group are again subdivided into three finer grades. There are thus in all nine qualitative grades—designated most conveniently by the letters dD, mD, fD (representing different grades of dullness), dM, mM, fM, and so on. The procedure has many obvious points of practical convenience ; and in several ways is closely similar to that adopted in this country by school medical officers who have no time or no preference for quantitative measurements in terms of mental age[1].

The metric scale of Binet and Simon has been in general use with those teachers and school medical officers in almost every

[1] An early account of the method is given in *L'Année Psychologique*, 1912, p. 327 : " L'Échelle Métrique de l'Intelligence de Binet-Simon Modifiée selon la Méthode Treves-Saffiotti." A critical summary of Saffiotti's latest volume, *La Misura dell'Intelligenza*, will be found in the *Eugenics Review*, VIII, iv. 1917, pp. 365–373.

part of the world who have concerned themselves with psychological tests. First intended mainly to pick out the defective, its scope has been progressively enlarged, until in America it has been applied, in one form or another to the measurement of intelligence among school children of every age and level and even to the measurement of intelligence among adults. It is to this day the favourite instrument for the diagnosis of mental deficiency in children. But for older and brighter individuals—such as candidates for Junior County Scholarships—it has now been almost entirely abandoned in favour of some other method—group tests, performance tests, or tests of reasoning and of higher mental powers.

30. (vi) *The De Sanctis Tests.* One other scale deserves description. A year after Binet published his first article on the measurement of intelligence, Professor de Sanctis, of the Laboratory of Experimental Psychology at the University of Rome, set forth in the same French journal a series of six tests, which he had worked out for grading mentally defective children[1]. Like Binet, de Sanctis had reached the notion of a step-like scale of problems, each harder than the last : but whereas Binet sought primarily to measure the amount of positive intelligence, de Sanctis was concerned chiefly to measure the degree of defect. De Sanctis, too, uses rather more apparatus than Binet—coloured balls, wooden cubes, pyramids and oblong blocks, a test-card representing squares, triangles, and rectangles, a screen for covering the apparatus, and a stop-watch for measuring the speed of the child's responses. The nature of the problems can be gauged by citing some of the seventeen test-questions given as applicable to feeble-minded children of almost any age. The examiner putting before the child the blocks of wood of various shape, and showing him a cube, says to him : " Pick out all the pieces that are like it." Or, again : " Look at this card : and

[1] *L'Année Psychol.*, (1906) pp. 70–84, " Types et Degrés d'Insuffisance Mentale." The Italian account was published in *Annali di Nevrol.* (Naples, 1906), " Tipi e gradi d'Insufficienza Mentale." A description in English is given in Whipple, *Manual of Mental & Physical Tests* (1st Ed. 1910, pp. 469–473. Omitted from 2nd edition). A leaflet containing a revised version of the tests (including minor modifications suggested by Madame Montessori and others) can be obtained from Professor de Sanctis (" *Reattivi De Sanctis* " *per la Valutazione dell'Insufficienza Mentale degli Anormali. Mod.* 1914).

point to the things that have the same shape as this piece of wood " (a cube). . . And later he asks : " Are big things heavier or lighter than small things?" . . . " When things are far away, do they look larger or smaller than things that are near ? "

Several attempts to standardise and evaluate the de Sanctis tests more precisely have been carried out during recent years[1]. In this country the chief investigation on their use is that undertaken at the Baldovan Institution for Feeble-minded Children by Dr. W. B. Drummond. From his results it would appear that the tests " afford a rapid and practical means of classifying the mentally defective," but are less suited for differentiating the defective from the normal, or for grading the intelligence of normal children even at the earlier ages. He writes that " the de Sanctis tests may be utilised as substitutes for some of the tests in the Binet Scale, but they cannot entirely take its place[2]." This judgment seems to express the general conclusion reached by the few psychologists who have tried the method.

X.—GROUP TESTS.

31. During the last six years, group tests, for long overshadowed by the Binet-Simon Scale, have come widely into practical employment. In the early days of mental testing the need for group tests was hardly felt. Children were tested, not in large numbers, but in exceptional instances—notably the comparatively rare cases suspected of mental deficiency. The want was first experienced, not for children but for adults, and not in the school, but in the army during the emergencies of the war. Among all the achievements of psychological examiners, the

[1] An early account of the tests, with some comments on their value for the diagnosis of mental deficiency will be found in Dr. Shrubsall's article on " The Examination of Mentally Defective Children," (*School Hygiene*, 1911, II. ii. pp. 609–612 ; a reprint of a report presented to the Committee of the British Association (Section L, Portsmouth, 1911) on " Mental and Physical Factors involved in Education)." The chief American papers on the subject are Goddard, " The Grading of Backward Children," reprinted from the *Training School Bulletin* (Vineland), 1908, and " Mental Development and Measurement of the Levels of Intelligence," *Journ. Educ. Psychol.* (1911), pp. 498–508, and L. Martin and " A Contribution to the Standardisation of the de Sanctis Tests " *Training School Bulletin*, (Vineland, 1916), XIII. pp. 93–110.

[2] " Observations on the de Sanctis Tests," *Brit. Journ. Psycohol.* (1920) X. ii. and iii. pp. 259–277.

testing of nearly 2,000,000 recruits for the American Army still remains the most remarkable[1]. In this examination the main object was twofold: to eliminate as rapidly as possible all who had not sufficient intelligence to be safely trusted with a rifle, and to discover all who possessed a sufficiently high ability to be immediately selected for training as commissioned or non-commissioned officers. Incidentally, however, the records and the results proved of great service in many other ways.

In order to test large numbers with the utmost speed, the oral and individual methods hitherto employed by most psychologists were supplanted almost entirely by the wholesale use of written tests administered to large groups simultaneously. The particular tests assembled for this purpose were in part those already used by investigators of the so-called higher mental processes, and in part a series of new test-problems devised expressly for the purpose in hand. So impressive and so successful were the results of this collective procedure, that, after the war was concluded, group-tests of a similar character were rapidly introduced into many American universities and schools for the examination of entrants. During the last few years there have been compiled, and issued from that country a large number of publications containing collections of group-tests of intelligence, the Terman Group Tests, and Otis Group Tests, the National Intelligence Scale, and many others[2].

The tests set at any one examination usually run to five or six in number, and each test may contain from twenty to fifty short questions. It is, indeed, one of the most suggestive discoveries of recent psychological work that, *for an efficient examination, it is far better to use a large number of short questions than a small number of long questions.* The tests and the questions are generally printed in the form of a booklet. One booklet is distributed to each candidate, who writes or marks his answers

---

[1] A complete and convenient account is to be found in Yoakum and Yerkes *Mental Tests in the American Army*. See also Appendix VIII.

[2] See tabular list given in the *Twenty-first Yearbook of the National Society for the Study of Education: Intelligence Tests and their Use.* (Public School Publishing Co., Bloomington, Illinois, 1922, pp. 93–113). The table gives the name, author, number and nature of tests, the ages for which they are applicable, and time required, the publisher, the price, and references (where available) to descriptive accounts.

on the pages according to instructions. To avoid difficulties in evaluating the various answers that might be written by the candidate himself, a number of alternative answers are often presented to him, and he is required simply to underline the correct reply. For example[1]:

1. How often do people { praise / fail / condemn } in others the very { faults / serious / virtues } they { are guilty of / pardoning. / perceiving. / themselves. }

In America it would probably be difficult to find, among all the larger and more progressive educational systems, any in which group tests of intelligence are not now being extensively applied. In Chicago, with a school population of nearly half a million, 50 per cent. of the children—chiefly those who are backward or advanced—are regularly tested by the principals or teachers. In Washington (D.C.), in New Rochelle (New York), in Cleveland (Ohio), in Denver (Colorado), in Kalamazoo (Michigan), all the schools are using the National Scale of Intelligence Tests or some similar set[2]. The tests are sometimes carried out by the ordinary teachers, but more usually by teachers specially trained, or by special officers and experts. In Washington, for example, "an expert psychologist newly appointed is to have a free hand in the classification of all pupils." For the most part it would seem that the tests are applied chiefly at the entering grades or at the time of promotion to junior and senior high schools[3]. The grading and the promotion are determined by the results, but not, of course, by the group-tests alone. The teacher's rating, the child's school record, and often the results of individual testing, are taken also into consideration.

The method of group-testing has been employed, upon a small and experimental scale, by school teachers and scholarship examiners in England. The Bradford Education Authority in 1919 adopted, for the purposes of junior scholarship examinations, a number of the written group tests first used in 1911 for an early research at Liverpool[4]. Two years later, at the request

[1] For further examples, see Appendix VIII.

[2] See Appendix III.

[3] See Appendix VI.

[4] *Annual Report of the Bradford Education Committee*, 1920; *Cf.* also Appendix II., p. 152.

of the Northumberland Education Authority, Professor Godfrey Thomson devised a set of group-tests for much the same object in what was one of the most notable experiments on the subject in this country[1]. It had been observed that nearly one-third of the schools in the county of Northumberland presented no candidates for the ordinary scholarship examinations in English and Mathematics. These schools were largely small schools in isolated rural districts, such as the Cheviots and the Dales; and it appeared possible that, from lack of home culture, of town life, and of teaching facilities, many of the best county pupils might be handicapped in essay-writing and arithmetic. Three thousand children were accordingly tested; and it was found that many of the most successful pupils resided in the remoter areas of the county. This early experiment was so successful that a group-test of intelligence has since been introduced on every occasion into the Northumberland examinations for such scholarships.

Similar difficulties have been encountered by other English education authorities; and have been dealt with experimentally by similar means. Dr. Ikin, Education Officer for Blackpool[2], has tested a group of one hundred scholarship candidates for Junior County Scholarships with five of the better-known group tests (the Terman, Otis, Northumberland, Simplex, and National Scale respectively). At Rugby, Mr. Vaughan has applied a set of group-tests, devised in London, first to certain selected forms and later to the whole of his school[3]. At Cheltenham Grammar School, Mr. Dobson has applied the same tests (with others) both to the entire school and to candidates for scholarship and entrance[3]. These are but a few of the more notable experiments upon these lines. Where the results have been statistically analysed, it is found that the calculated correlations show a close correspondence with the results of independent scholarship examinations or of independent personal judgments. Where the tests and the scholarship examinations disagree, subsequent study of the children shows that the test has often revealed inborn ability which the scholastic examination failed to detect, owing to the child's lack of opportunity, at school or at home, for

[1] *Brit. J. Psychol.* XII. 201. 1922. *Cf.* Appendix II., pp. 153–155.

[2] See *The Times Educational Supplement*, Oct. 6, 1923, p. 444, "Group Intelligence Tests." *Cf.* Appendix II., pp. 149–151.

[3] See Appendix II.

acquiring the necessary knowledge. None of the investigators, however, has as yet claimed that intelligence tests can do more than supplement written examinations of the ordinary scholastic type. Before a child can be admitted to a Secondary School he must possess a certain minimum of educational knowledge ; and this is to be gauged, not by a test of mental capacity, but by a test of scholastic acquirements.

Group-tests have also been introduced into examinations for adults. At the London Day Training College, psychological tests have been employed for the last two years as a supplementary means of selecting candidates for the four year course of training for teachers[1]. At Bedford College, since 1921, the Psychological Department has tested incoming students of different faculties with tests designed to measure " Arts " ability, " Science " ability, and " general ability " or intelligence At University College, the staff of the Psychological Laboratory has similarly employed group-tests to test the intelligence of freshmen who have volunteered to sit for the examination. These experiments are still too recent for any adequate comparison to be made with the subsequent academic careers of the candidates. In this country, however, the most extensive use of such group tests has been the introduction of a psychological test-paper into the competitive examination for clerical posts in the Civil Service in 1920 and the following years. Nearly 40,000 candidates have been so tested ; and an analysis of the published mark-lists shows that the psychological test correlates with the general results more closely than any other single paper.

[1] The candidates are young men and women between the ages of 17 and 19. A preliminary experiment was made in 1922, and the results were sufficiently useful to warrant an extension of the method. In 1923 all applicants took a half-hour's test-paper, similar to that devised for the Civil Service competition (used also experimentally at the Bristol University Department of Education and at several other Training Colleges). The evidence relied upon in estimating a candidate's merits is thus derived from three sources : first, the report of his Secondary School ; secondly, the impression produced by him during an interview—the interview itself being partly standardised in form ; and, thirdly, the results of the psychological test. The test-results are valuable, not only as an independent source of information about individual candidates, but also as a means of equating the varying standards of assessment implied in the reports from the different Secondary Schools. The estimates based independently upon each of these three methods will ultimately be compared with subsequent progress of the candidates in their academic and professional work.

XI.—PERFORMANCE TESTS.

32. Both the Binet-Simon Scale, and the majority of the group-tests in present use, are predominantly linguistic in character. For the most part they consist of verbal questions; and call for verbal answers. As a rule, this is an advantage; but in exceptional cases it may prove a serious drawback. In America, the large number of illiterates, and of alien children and adults who speak no English, gravely limits the use both of the Binet-Simon Scale and of the ordinary written tests. Occasionally, too, even in this country, the examining officer encounters a child who for various reasons is heavily handicapped for any test of a verbal type. Such a child, whether from nervousness or from some defect in hearing or speech, may be utterly unable to do himself justice in an examination consisting of nothing but oral questions and answers. Owing to the low culture of his home, or a lack of regular attendance at an ordinary school, he may possess not even that bare minimum of knowledge which the Binet-Simon Scale assumes; sometimes even an older and more intelligent child is unable to read, write, compute or recognise the common coins. The profound influence of these shortcomings has been clearly demonstrated in a recent research by Mr. Hugh Gordon upon gipsy and canal-boat children[1]. Such cases are exceptional, but they call for constant vigilance in routine examinations.

To surmount this defect, several psychologists have sought to construct supplementary tests of a practical rather than a verbal character, known as Performance Tests. Instead of answering by word of mouth the child is required to do something, or to make something, with his hands. He is asked to fit together pieces of cardboard or wood so as to make certain shapes or pictures; to build up a large cube out of little cubes; to make

---

[1] *Mental and Scholastic Tests Among Retarded Children*, Board of Education Reports, 1923. Mr. Gordon's statement of the limitations of the Binet tests—based upon the American version—are quite consonant with the opinion generally held in this country. In the press, however, his statements have been carried beyond the cautious point to which he himself took them; and have been treated as criticism of the general use of such tests for ordinary purposes with children of ordinary schooling. For a reply to these too sweeping deductions, see the review by Dr. P. B. Ballard, "Mental and Scholastic Tests among Retarded Children," *Forum* I., iii. (1923), p. 250.

various movements in imitation of the examiner; or to number geometrical figures according to a key. Nearly all such tests need special apparatus to be obtained from manufacturers of scientific materials.

Of these Performance Tests the oldest example and the commonest type is the form-board devised by Edouard Séguin (1846). It consists of ten wooden blocks of various shapes—a square, a triangle, a circle, a star, or a Maltese cross; each block has to be fitted into a hole of similar design; the time and child's procedure is recorded. Séguin intended his form-board for training rather than testing; but in this country it has long been used as a test for the mentally defective. By altering the number, shapes, and subdivisions of the blocks, form-boards increasing in difficulty have been devised. A different type of non-linguistic test, used with success in an early research in England and since applied in various forms and upon an extensive scale in America, consists in the reconstruction of dissected pictures. This test, as we have seen, was suggested by Ebbinghaus' "combination" theory of intelligence, embodied, however, in concrete instead of verbal material. In its simplest form, a suitable picture-postcard was cut up into eight or twelve small rectangles, and the child had to re-combine the fragments into a whole, with or without an intact copy before him—after the fashion of a jig-saw puzzle[1]. Yet another early test, also used in the first instance for testing the feeble-minded, is the age-scale of graded mazes devised by Porteus (1915)[2]. A recent collection of fifteen of the better tests of the performance type has been made by Pintner and Paterson, who have drawn up standard instructions and published norms of achievement for children of various ages[3].

---

[1] Experimental Tests of Higher Mental Processes and their Relations to General Intelligence." *Journ. Exp. Ped.*, 1911, I., ii., p. 102. For later work with tests of this type, see Pintner and Anderson, *loc. cit. inf.*

[2] *J. Exp. Ped.* (1915), III., ii., pp. 127–135. "Motor Intellectual Tests for Mental Defectives." The necessary materials will be found reprinted in *A Handbook of Tests for Use in Schools* (P.S. King & Son, 1923).

[3] *A Scale of Performance Tests* (D. Appleton & Co., 1917). Another collection is described by Dearborn, "Form-Board and Construction-Tests of Mental Ability." *J. Educ. Psych.*, VII (1916) pp. 445 *et seq.* *Cf. id.* "A Series of Performance Tests of Intelligence," *Harvard Mon. Educ.* I., iv., 1923.

Performance Tests have been employed far more widely in the United States than elsewhere. The chief uses to which they have been put are the following :—

(i) To test the intelligence of the deaf. It was, indeed, for the purpose of classifying school-children who were deaf or hard of hearing, that the first scale of Performance Tests was devised by Pintner and Paterson[1].

(ii) To test the intelligence of speech-defectives. Persons who have an impediment, whether physical or nervous, in their speaking, generally fail to give a just impression of their powers in the customary verbal tests.

(iii) To test the intelligence of foreigners who speak no English, and are often illiterate. Such tests are now extensively used with immigrants at Ellis Island, New York ; and were employed for a similar purpose in the United States Army during the war[2].

(iv) To test children and young persons, who, from lack of schooling, are unlikely to do themselves justice in verbal and literary tests. For this purpose Performance Tests have been largely used in testing delinquents, truants, defectives, and children from ignorant homes.

(v) To test children and young persons, who, on the ground of special linguistic facility (so-called " verbalists ")[3], are likely to show with too great an advantage in conversational tests of the Binet-Simon type.

(vi) To test the intelligence of young persons for vocational guidance, where practical rather than intellectual ability is chiefly required.

Experiments with such tests have lately been carried out upon school children in London by an investigator from the psychological laboratory of Bedford College, London, and, independently, by a second investigator trained in America and

---

[1] See Pintner and Paterson, *loc. cit. sup.* ; also *id.*, " Learning Tests with Deaf Children," *Psych. Rev. Mon.*, XX., iii., p. 87 (1915), *Cf. id.*, " A Class Test with Deaf Children," *J. Educ. Psych.* VI. (1915), p. 591.

[2] Knox, H. A. " A Scale based on the Work at Ellis Island for estimating Mental Defect." *Journal of the American Medical Association*, 1914. See also Yoakum and Yerkes, *Army Mental Tests.*

[3] See Healy, W., *The Individual Delinquent*, pp. 473 *et seq.* Also Healy, W. & Fernald G., *Tests for Practical Mental Classification* (1911), where some of the first and best-known performance tests are described.

working under the Industrial Fatigue Research Board—a department of the Medical Research Council. The results[1] show that tests of this kind are well adapted for measuring the practical intelligence of children who with a more linguistic type of test fail to reveal their true ability, such as the canal-boat children alluded to above. Further, these tests not only arouse immediate interest in young people who show a distaste for an examination of an ordinary scholastic type, but also incidentally bring to light many important qualities of temperament and personal outlook. It is clear, however, that both the methods and the results need further standardisation before they can be applied in this country with success.

XII.—STANDARDISED TESTS OF SCHOLASTIC ATTAINMENT.

33. In a history confined primarily to psychological testing the development of tests, educational attainments need only be touched upon very briefly.

34. *The Standards of Former Codes of the Education Department.* The conception of standard tests, based on age-performance and measuring educational capacity and attainment, is by no means a new idea, due wholly to recent psychological research. The Report of the Newcastle Commission on the state of popular education in England issued in 1861 a recommendation that a Grant should be paid in respect of every child, who, having attended the Elementary School in the year preceding the day of examination, had passed an examination in Reading, Writing, and Arithmetic[2]. As a result of this recommendation, a provision was introduced by the Committee of Council on Education into the Revised Code for 1862[3], stipulating that every scholar for whom Grants were claimed must be examined according to one of six standards (each briefly described by the Code) in the subjects specified. The ultimate result of this regulation was

[1] To be published shortly.

[2] Report of Royal Commission on the state of popular education in England (1861), Vol. I., p. 545, Recommendation 6.

[3] Revised Code for 1862, Arts. 40, 46–48. The issue of the Revised Code for 1861, Arts. 43 and 44, had provided for four groups of children of the following ages, 3 to 7, 7 to 9, 9 to 11, 11 and upwards, with standardised ranges of attainment for each group in reading, writing and arithmetic.

the organisation of elementary schools on a basis of annual promotion. Each class in the senior department corresponded to an age group; and the whole series of classes were numbered standards I. to VI.[1], roughly corresponding to the ages 7 to 12. The test-cards set by the inspectors were in some ways strikingly similar to the group-tests of scholastic attainments drawn up for successive years by psychologists of the present day.

From the first there was much opposition to these arrangements. The objections of the teachers, however, were directed not so much against the method of testing as against the principle of "payment by results." The merits and defects of the standards then formulated, obvious as they now seem in the light of recent research, are not entirely uninstructive. These standards were based, not upon an experimental enquiry into what children of a given age actually knew, but upon an *a priori* notion of what they ought to know: they largely ignored the wide range of individual capacity; and the detailed formulations for the several ages were not always precise or appropriate[2]. In the course of thirty years the first strict conditions were gradually relaxed, more and more freedom of classification being given, the tests made more and more elastic, and the examinations being taken by sample only. From about 1892 the system of examination by standards began to fall into disuse; and it was finally abandoned by the Board of Education about the beginning of the present century except for a few special purposes, such as examining candidates for Labour certificates.

35. *American Tests.* Outside Great Britain the first application of the newer methods of scientific measurement directly to the results of school teaching was made in 1898 by an American, Dr. J. M. Rice, in an investigation upon spelling[3]. The protests of American educational experts and teachers against stereotyped

---

[1] A Seventh Standard was added about 1882.

[2] A conspicuous instance is the statement that the child should be able to spell words from the same books as he used for reading. Recent work shows very clearly that to spell a given word is much harder than to recognise it when spelt already in print. As a rule a child can read most words a year earlier than it can spell them. In general, almost the whole of the first requirements proved to be too hard and stringent by the equivalent of about a year; and later the original requirements for Standard I were fixed for Standard II, and so on.

[3] "The Futility of the Spelling Grind," *Forum*, XXIII., pp. 163–172, and 409–419.

methods of examination for attainment were even more vehement than the earlier protests in England. But five years later a committee on school efficiency was appointed at the Philadelphia meeting of the Department of Superintendence; superintendents themselves were rapidly converted; and the custom of school surveys, including investigations of educational attainments as well as of attendance, costs and equipment, developed with great speed, if not always with equal prudence.

The earliest set of educational tests, however, standardised upon lines resembling those adopted in the measurement of intelligence, emanated once more from France, and was the work of three Parisian psychologists. In conjunction with Monsieur V. Vaney, Binet and Simon attempted, besides their scale for natural intelligence, what they styled a "barometer of instruction" —a set of graded exercises in reading, spelling and arithmetic[1]. The tests were compiled upon a rough and ready plan; and were of necessity adapted only to French requirements.

In America, under the enterprising lead of Professor E. L. Thorndike, a number of test-scales have been issued from time to time during the last twelve years, based upon an elaborate statistical analysis, and dealing with the measurement of the various subjects of the Elementary School curriculum[2]. Of these the earliest appears to have been Thorndike's scale for estimating quality of handwriting[3]. For this purpose, however, the scale which has been the more widely used is that constructed by Ayres, based, not like the Thorndike scale upon the general merit of the script, but simply upon its legibility. Since the early experiments of Rice, the efficiency of spelling has been the subject of several tests, notably by Ayres[4] and Buckingham[5]. Reading has been the subject of innumerable

---

[1] *The Development of Intelligence* (1905), (Kite's translation) p. 70; see also *Mentally Defective Children*, (Drummond's translation) p. 54.

[2] A compact description of American tests of scholastic attainment will be found in Wilson & Hoke, *How to Measure* (Macmillan & Co., 1920). A reference to the bibliography in Appendix VII., however, will show that American publications on this subject are almost innumerable.

[3] Thorndike, "Handwriting," *Teachers College Record*, March, 1910. Ayres, "A Scale for Measuring the Quality of Handwriting of School Children." *Russell Sage Foundation*, Bulletin No. 113.

[4] *A Measuring Scale for Ability in Spelling.*

[5] *Spelling Ability: Its Measurement and Distribution.* Teachers' College, Columbia.

tests. Among these perhaps the most notable are those of Thorndike, Kelley and Munroe[1]. In arithmetic the results of the long researches carried out by Courtis deserve especial mention[2]. For drawing and for composition scales once more have been drawn up by Thorndike[3]. More recently tests have been devised for history and for geography as well as for algebra, Latin, physics and chemistry.

Apart from differences in idiom and in values for money, weights, and measures, the chief disadvantage of the American scales, lies in the custom of compiling averages and norms for school-classes or "grades" instead of for age-groups. Such figures cannot, therefore, be adopted, as they stand, for English children or in English schools.

36. *English Tests.* In this country the pioneers have been Dr. Ballard, Mr. Winch, and the late Professor J. A. Green. Dr. Ballard was early in the field with simple and effective tests of arithmetic[4] and reading[5]; he has lately collected a number of test-scales, compiled by himself and others, in his books on *Mental Tests* and *The New Examiner.* Both as editor of the only periodical in this country dealing solely with educational psychology[6], and as secretary of a Committee appointed by the British Association to enquire into Mental Factors in Education, Professor Green, together with his research-students, was among the first to introduce and to investigate, not only the Binet Scale itself, but also experimental tests of particular school subjects, in their application to English schools. Mr. Winch has employed original tests of arithmetic and other school subjects to investigate memory, fatigue, reasoning, and

---

[1] Thorndike, E. L., "*Improved Scales for Work Knowledge or Visual Vocabulary.*" Kelley, F. J., "*The Kansas Silent Reading Tests.*" Munroe, W. S., "*Standardised Silent Reading Test.*"

[2] "Measurement of Growth and Efficiency in Arithmetic," *Elementary School Teacher*, X., and later numbers.

[3] "The Measurement of Achievement in Drawing." *Teachers College Record*, 1913. "Thorndike Extension of the Hillegas Scale for the Measurement of Quality in English Composition." 1912.

[4] *J. Exp. Ped.* 1914, II. p. 396.

[5] *J. Exp. Ped.* 1915, III. p. 153.

[6] *The Journal of Experimental Pedagogy*, now re-named *The Forum of Education*, and edited by Professor Valentine, of Birmingham.

the transfer of training in children[1]. More recently, the London County Council has issued a set of standardised tests not only for native intelligence but also for attainments in the chief subjects of the elementary curriculum—such as reading, spelling, arithmetic, writing, drawing and composition[2]; and has published the results of an educational survey of a representative borough, carried out with the assistance of teachers by the application of such tests[3]. Of these later tests the majority are arranged upon an age-basis: there are, for instance, for each successive year, ten words which the average child can read, ten words which he can spell, ten sums in mental, mechanical, and problem arithmetic which he can work, at the age assigned. It is then possible quickly and easily to compute an "educational age" for every individual scholar in all the chief subjects of the elementary curriculum[4].

XIII.—TESTS OF VOCATIONAL APTITUDE.

37. The other important branch of applied psychological testing, namely, the testing of vocational aptitudes, bears upon the work of the schools only in an indirect fashion and in certain limited fields. Consequently, despite its rapid and remarkable expansion during the last five years, its history need here be but briefly recounted.

By vocational psychology is understood the discovery and measurement, by scientific methods, of those special mental qualities in virtue of which a particular individual is naturally adapted for one occupation rather than for another. It is generally regarded as embracing two main divisions, termed respectively vocational guidance and vocational selection.

---

[1] *J. Educ. Psychol.*, 1910, I, i. and ii., 1916, VII. ii. *Child Study*, 1913 VI *et seq. Brit. J. Psychol.*, 1911, IV, ii. 1914, VII. ii. *J. Exp. Ped.*, 1913, II. ii, 1921 VI. iii. For the meaning o the phrase "transfer of training," see above, section 11, paragraph 2.

[2] *Three Memoranda on Mental and Scholastic Tests*, 1921. The test materials contained in this volume have been re-issued in the form of a *Handbook of Tests for Use in Schools*.

[3] *Three Preliminary Memoranda on the Distribution and Relation of Educational Abilities*, 1917.

[4] For examples of such tests, see Appendix VIII.

Vocational guidance aims at finding the best occupation for a given person; vocational selection aims at finding the best person for a given occupation. The former falls within the scope of the teacher, so far as he has to prepare and recommend his pupils for whatever trade or profession may suit each best when due to leave the school; the latter falls within his scope, so far as he has to select particular boys or girls for Trade Schools, Art Schools, Apprentice Schools, Central Schools with an industrial or commercial bias, or similar institutions such as prepare their scholars for definite forms of employment.

In neither instance are the aims of psychological guidance and selection wholly new. For many years the teacher who has taken a personal interest in the children under his charge, has habitually aided them with advice upon the kinds of employment best suited to the special aptitudes of each. The numerous professional examinations—those which the lawyer, the doctor, the schoolmaster, the engineer, the accountant, the civil servant, must pass before they are qualified to take up their chosen work—are, in essence, vocational tests. And the well-known Board of Trade tests for colour-blindness are strictly psychological tests, worked out in the laboratory by psychological methods, for what is definitely a problem of vocational selection.

But with one or two rare exceptions of this type, the tests and devices used until recently, both for selection and for guidance, have been eminently unscientific. In its scientific form, the history of vocational psychology is usually dated from two American experiments carried out during the first decade of the present century.

38. *Vocational Guidance.* The beginnings of psychological "guidance" upon systematic and scientific lines are generally traced to the experiments of the late Mr. Frank Parsons in Boston, U.S.A. Parsons began by arranging conferences with all the boys of his neighbourhood who were to leave the Elementary Schools at the end of each year. His recommendations were based upon answers to a psychological questionnaire rather than upon performances in psychological tests. But out of these informal discussions there grew up a permanent office which was opened in 1908, where all Boston boys and girls were able to come for counsel and advice upon the choice of a vocation[1].

[1] Frank Parsons, *Choosing a Vocation*, Boston, 1909.

The Boston Vocation Bureau, thus established, rapidly stimulated a large number of American cities to come forward with similar plans. Educationists have been especially attracted by the movement. Most American High Schools now have their vocational adviser; and a large number of educational systems have offices for vocational guidance. Tests of general intelligence and educational attainments are freely used; and tests for particular occupations have been introduced more recently, borrowed from the increasing array of specific trade-tests worked out primarily by those engaged not in guidance but in selection.

39. *Vocational Selection.* The beginnings of psychological "selection" are usually traced to an earlier experiment by Mr. F. W. Taylor, an American engineer[1]. In the factory of a Rolling Machine Company, in Massachusetts, Taylor endeavoured to select the best girls for the work of inspecting bicycle-balls by means of a test of reaction-time. Some of the girls, already engaged as inspectors, were found to show a slow reaction-time, and were dismissed—although this involved the dismissal of "many of the most intelligent, hardest working, and trustworthy girls." The broad result was that, after selection, thirty-five girls did, in a shorter working day, as much work as one hundred and twenty had done before selection.

Neither the experiment of Parsons nor that of Taylor, was carried out upon what would now be considered strictly scientific principles. Their work, however, aroused active interest; and their writings, and those of other "efficiency-engineers," stimulated Professor Hugo Münsterberg, at that time head of the psychological laboratory in Harvard University, to publish a systematic account of the possibilities of industrial psychology, dealing in special detail, among other topics, with test-methods for selecting the most suitable work and the most suitable workmen[2]. Münsterberg himself carried out several experiments on vocational selection. He devised ingenious schemes of testing for the telephone service, for the electric railway service, and for navigating officers of a large ship's company. His oft-cited experiment on the selection of tram-drivers appears to have

[1] *Principles of Scientific Management,* (Harper Bros. 1911.) The experiment was first described in 1903.

[2] *Psychology and Efficiency.* (Constable & Co.) 1913.

been the first real experiment in vocational selection by genuinely psychological tests[1].

40. *Vocational Testing during the Great War.* Starting in this somewhat sporadic and haphazard way, the usefulness of vocational testing was developed and demonstrated during the war. In England, under the Air Board and under the Admiralty, tests were devised and carried out by laboratory psychologists for air-pilots, aeronautical observers, hydrophone operators, and for many other military and naval tasks needing special capacities of sense-perception or special degrees of intelligence. In America trade-testing was carried out in even greater detail. Skilled or efficient men were required immediately for over four hundred separate occupations. There was no time to train them ; or to try them out by a period of probationary engagement. Misfits might mean a grave disaster. Hence, various trade-tests—based upon the usual principles, and of the oral, pictorial, " performance," and " written-group " types—were rapidly and successfully devised[2]. The work was undertaken by a Committee of the U.S.A. War Department (Committee on Classification of Personnel) operating through an Army Trade Test Division established in three separate centres ; and was largely guided in its beginnings by Professor E. L. Thorndike (who had already

---

[1] Munsterberg's method of selecting the best telephone operators was based upon a collection of eight tests : (1) immediate memory ; (2) logical memory ; (3) speed of movement ; (4) accuracy of movement ; (5) speed of association ; (6) spacial judgment ; (7) card sorting ; (8) cancelling certain letters from a page of print.

The method for testing the drivers of electric street cars involved special apparatus. This was a machine, exhibiting at a regular speed, through a slot or window, a card which gave a diagrammatic representation of a street and of different types of moving objects. Two parallel lines running through the centre of a card lengthwise represented the tram lines. The whole card was divided into small squares. In certain of these squares the figures 1, 2, and 3 were placed, 1 representing a slow-moving pedestrian, 2 a horse going twice as fast as a pedestrian, and 3 a motor-driven vehicle going three times as fast as a pedestrian. Objects moving across the track were coloured red ; those moving parallel with it, black. Each square along the track was lettered from A to Z. The subject was required to name as fast as possible the squares on the track threatened (as it were) by the red figures.

[2] See J. Crosby Chapman, *Trade Tests : The Scientific Measurement of Trade Proficiency* (George Harrap & Co.) 1922.

done so much in establishing other branches of psychological testing) in conjunction with Colonel W. D. Scott and Dr. W. V. Bingham. After the war the same methods were naturally continued by the larger firms, and by such bodies as the Carnegie Institute of Technology and the American Civil Service Commission[1].

The movement thus begun has rapidly spread. France, Belgium, Holland, Spain, Switzerland, Germany—indeed, most of the civilised countries of the world, now possess institutes for vocational guidance, in which trained psychologists, working in close contact, not only with business firms and psychological laboratories, but also with school teachers and education authorities, carry out vocational tests and offer vocational advice.

41. *Vocational Guidance among English School Children.* In this country two main bodies have carried out investigations upon vocational guidance and selection. The Industrial Fatigue Research Board, a branch of the Medical Research Council, was established during the war to carry out, by physiological and psychological methods, scientific enquiries upon industrial efficiency. It includes a psychological committee; and, since the war, has carried out several enquiries upon vocational selection; and has published a useful review of the literature of vocational guidance[2].

The National Institute of Industrial Psychology, founded in 1921, includes among its principal aims "the more efficient guidance of children in taking up their life's work"; and has recently established a vocational section. Its investigators have already carried out a number of experiments upon selection in schools, offices, factories, and firms of different types; and have constructed and published group-tests for estimating "intelligence" in older children and younger adults, and specialised tests for estimating aptitude for particular professions and trades[3].

---

[1] See more particularly the *Fortieth Annual Report of the U.S. Civil Service Commission* (1923), especially "Report of Research Section," pp. li.–xcix.

[2] *Reports of the Industrial Fatigue Research Board, No.* 12. "Vocational Guidance." By Professor B. Muscio. (H.M. Stationery Office, 1921), 1*s.* net.

[3] The results of these investigations are recorded from time to time in the *Journal of the National Institute of Industrial Psychology*. See specimen tests contained in Appendix VIII.

Under the Education (Choice of Employment) Act, 1910, many local education authorities have drawn up schemes for placing children in suitable employment when they leave school. These recommendations are based largely upon the reports of the child's head teacher, usually entered upon a so-called "school-leaving form." The report includes general observations as to the child's ability, character, and conduct, a statement of the sort of employment which he desires and the teacher recommends, together with notes extracted from the records of the medical officer's last inspection. In London during the past eighteen months an investigation has been carried out in certain schools in a selected borough to determine how far these reports and recommendations can be made still more effective, by basing them upon a scheme of intensive psychological testing. A survey has been carried out of recent placements within the area chosen; the commonest occupations so secured have been analysed; psychological tests have been administered for general intelligence, for special attainments, and for the qualities needed for these commoner trades; and the children are being followed up to see how far the recommendations thus deduced have proved sound and beneficial. The experiment has been carried out as a joint research by investigators under the London County Council, the Industrial Fatigue Research Board, and the National Institute of Industrial Psychology.

42. *Vocational Selection for Trade Schools and Apprentice Schools.* The establishment of Trade Schools and similar institutions, and the award of trade scholarships, has made tests and examinations for trade aptitudes a definite element in the work of education authorities. As a rule, such examinations are partly based upon subjects of the traditional type (English—including history, geography and composition—and arithmetic); but also include more technical or practical subjects—such as drawing (freehand, geometrical, object, and nature drawing), woodwork or metal work for boys, and cookery or needlework for girls, according to the type of scholarship to be awarded, and on these special weight may be laid. In London, by the initiative both of the Principals of three or four Trade Schools and polytechnics, and of investigators under the National Institute of Industrial Psychology, experiments have recently been made upon the possibility of adding to the entrance examinations vocational tests upon lines more definitely psychological. In the

apprentice-schools attached to two large engineering firms at Manchester, trained psychologists have also been enquiring into how far existing methods of selection can be improved by the introduction of more scientific modes of testing. With an equal measure of success, experiments have also been carried out with tests for the selection of dressmakers, shorthand-typists, printers, and various other trades. Several large English factories have appointed their own works' psychologists, whose duties are largely concerned with the selection of young people for their Apprentice Schools and workshops. The whole enquiry is still in an experimental stage, but positive results, of benefit to the firms and to their employees, to the Trade Schools and to their pupils, have already been obtained[1].

XIV.—TESTS OF TEMPERAMENT AND CHARACTER.

43. Among the many factors that determine a child's educational progress, temperament and will are not less important then intelligence or knowledge. Accordingly, in the study of individual pupils, it is always desirable, though by no means always possible, to form an opinion upon the moral and emotional qualities of the child, as well as to test his intellectual capacities and attainments. This, indeed, is the very latest field into which psychological testing has penetrated.

44. *The Testing of Neurotic and Delinquent Children.* The analysis of character has been found particularly needful with those who suffer from moral or emotional disabilities, namely the delinquent and the nervous or neurotic. Of late years there

[1] A brief account of the history and aims of vocational guidance will be found in Professor Claparède's little pamphlet on *Problems and Methods of Vocational Guidance*, (International Labour Office, Geneva, 1922); and some concrete suggestions on the possibility of vocational diagnosis in schools and among children will be found in the *Lectures on Industrial Administration*, edited by Professor Muscio (Pitman & Sons, 1920.)

has been an increasing tendency, both in this country and abroad, for neurotic children and juvenile delinquents to be referred for examination to a psychologist or a psychological clinic, wherever the facilities exist[1]; and, in all such cases, the measurement of intelligence and the study of temperament must go hand in hand. Neither is adequate without the other.

45. *The Influence of Emotional and Moral Factors on the Testing of Intelligence.* But even with the normal, healthy child, emotional and moral qualities are apt to disturb the results of psychological testing in whatever form. In a group test of intelligence the lazy child may fail to exert his utmost powers. In an oral examination, as in the Binet tests, the shy or timid child may become suddenly confused, or nervously apprehensive, or altogether paralysed and mute. Hence, the testing psychologist must remain ever alert, lest, when he thinks he is testing intelligence, his final results may be vitiated by the irrelevant intrusion of excitement or emotion. And to overcome or counteract these tendencies numerous rules of procedure and devices of technique have of late been progressively contrived[2].

46. *Tests of Temperament and Emotion.* Increased efforts have recently been made to measure temperament and character by methods more direct and exact, by tests similar to those which have been used for the precise estimation of "intelligence." But, although the idea of temperamental testing is almost as old

---

[1] In Birmingham two psychological experts to assist the Justices have recently been appointed, one for adults and one for children. In London, teachers, magistrates, and organisers of children's care may refer neurotic or delinquent children for examination by the Council's psychologist; and an increasing number of such cases are now dealt with by such means. The Glasgow Educational Authority, as well as that of London, has now a psychologist attached to its staff, whose duties are primarily concerned with the testing of subnormal children (particularly defectives) and with the training of teachers in suitable methods. In America most large cities have their psychological clinic attached to the Courts, the Universities, or the local administrative school authority. See Appendix III.

[2] A suggestive paper, dealing with some of these recurrent difficulties, is that of Augusta Bronner, on "Attitude as it affects the Performance of Tests." *Psych. Rev.*, 1916, XXIII, iv.

as that of intellectual testing, it has had quite a different career[1]. Early observations and experiments upon the estimation of character and temperament were carried out by both Galton and Binet. Binet, in particular, besides his scale for intelligence, carried out many ingenious researches upon the measurement of conscientiousness, suggestibility, and fidelity of description and report[2]. But the first to approach the problem with an adequate statistical procedure was Dr. Naomi Norsworthy. The procedure itself was borrowed from that used by Professor Cattell in a biographical study of American men of science. From her analysis of the various estimates of a teacher's personal character —estimates made independently by half-a-dozen students— Miss Norsworthy concluded that " it would seem possible, by the use of some such method as the foregoing " (measurement by relative position or ranking) " carried out on a much wider scale, to justify a list of character-traits, numerical estimates of which by competent people would be both reliable and significant[3]."

Of all methods of investigating emotional tendencies by experimental methods, the oldest, the best-known and the most widely used are the methods of associative reaction and of the so-called psychogalvanic reflex. The two are frequently employed in conjunction.

There is a story which tells how an amateur detective, travelling one day by train in a full compartment, offered to guess the professions of his fellow passengers if each would give him, first of all, two meanings for a single word. The word was " box." Everyone began by saying that a box was a receptacle, of some firm material, and almost any size and shape—or some such phrase. But their second definitions differed altogether; and, according to the tale, at once disclosed some personal or professional interest. A lady, whose dress and complexion alone betrayed her calling, said a box was " the best place in a theatre."

---

[1] A good summary of the literature, with a detailed bibliography, will be found in Cady's article on " The Psychology and Pathology of Personality: A Summary of Test-Problems," *J. Delinq.*, VII., 225 (1922).

[2] *L'Étude Experimentale de l'Intelligence*, Paris, 1903.

[3] " The Validity of Judgments of Character." *Essays in Honour of William James.* 1908, pp. 553–567.

The postman said it was "something you received at Christmas." The schoolboy thought it was "something you received upon the ear." A pugilist said it meant "fighting with the fist." A baseball enthusiast explained that it was "the square where the pitcher stood." A soldier described a sentrybox; a Scottish landowner a shooting box; and an engineer started talking of gear-boxes and axle-boxes.

It is upon this principle—the varying directions taken by our spontaneous processes of thought—that the test of the associative-reaction primarily depends.

In its scientific form the experiment was first suggested by Galton fifty years ago; in his own case he found it gave a surprising insight into the ideas and contents of his mind[1]. A test of free association, of the continuous type, was later inserted by Binet into his scale of tests for intelligence; Binet asks the child to say in succession as many disconnected words as he can, in the space of three minutes. Later still, in the present century, the possibilities of such a test for detecting more emotional interests, whether conscious or unconscious, and particularly for discovering so-called repressed complexes, were developed by the psycho-analytic school, most of all by Dr. C. G. Jung and his pupils[2]. The method devised by these last investigators, and now in current use, is roughly as follows. A list of words is first prepared, the majority calculated to arouse some emotional experience or memory; the words are called out, one by one, to the subject, and he is required to answer as quickly as possible with the first word that comes into his mind: *e.g.*, "Box?" "*Theatre*"; "Paper?" "*Fire*"; "Wish?" "*Drink*"; "Fear?" "*Mouse*"; "Teacher?" "*Preacher*"; and so on. His answers are then compared with a standard collection already compiled from a thousand tested persons; the comparative frequency of each reply is noted; and the time each takes is measured, in fractions of a second, with a stop-watch or a chronoscope.

The so-called psycho-galvanic reaction was discovered accidentally by a Swiss engineer, E. K. Müller; and the experiment

---

[1] *Inquiries into Human Faculty* (1883) pp. 133–146. "Psychometric experiments."

[2] *Studies in Word Association.* (1918: a reprinted collection of earlier papers.)

was studied psychologically by Jung and his colleagues at Zurich[1]. In this country the method was first tested and confirmed in the psychological laboratory at Liverpool University; and has since been the subject of numerous researches[2].

The discoverer, who happened to be holding the wires of a galvanometer, noticed the curious fact that, whenever he experienced an emotion, the resistance offered by his body to the passage of the current seemed to be momentarily lowered. It was afterwards found that, if electrodes be fastened to a person's hand, the deflections of the galvanometer show instantly whether his thoughts—excited it may be either by an inward reflection or by an outward stimulus such as a name or a pistol-shot—are taking an emotional turn. With one and the same person, the degree of the deflection corresponds roughly to the intensity of his feeling; hence his chief emotional interests can easily be explored. Between different persons, comparisons are more difficult to draw, but, after certain allowances have been made for differences in their body-resistance, the intensity or frequency of the deflections (as the most recent experiments suggest) will probably provide us, after further research, with one of the most trustworthy methods for measuring emotional susceptibility and temperament.

---

[1] Jung, *Loc. cit.*, chapter XII. So early as 1888, however, the effect of psychical processes on resistance to electrical conduction had been studied by French scientists (Féré, *Comptes rendus de la Société de Biologie*, 1888, p. 217 *et seq.*; Vigouroux, *Le Progrés Médical*, 1888, sem. i, pp. 45 and 86). The Swiss investigations appear to have started from an independent discovery of the phenomenon, as described above.

The cause of the phenomenon is still not accurately known. It has been supposed that every emotion is accompanied by a slight reflex increase in the perspiration of the skin; the skin being the chief seat of the resistance to the current, the moistening of the parts in contact with the electrodes would thus temporarily lower the resistance offered.

[2] The fullest account in English is to be found in Whately Smith's book, *The Measurement of Emotion*, 1922. The most careful research carried out in this country is that of Dr. E. Prideaux, who investigated the method in the Cambridge Psychological Laboratory for the Medical Research Council ("Expression of Emotion as shown by the Psycho-Galvanic Reflex," *Brit. J. Med. Psych.* II., i., (1921) pp. 23–46).

47. *Tests of Character and Morality.* Moral tests have also been investigated by similar means[1]. In this country, however, the few who have investigated the possibility of tests for character have preferred an indirect to a direct technique. The moral test is, as it were, disguised as of a test of intelligence or information. A device, full of possibilities in this direction, is the optional question-paper. Every teacher knows how, in an examination on languages or mathematics, the routine-worker chooses the mechanical question, while the more enterprising select the problems and the riders; the cautious keep to the prepared texts, while the adventurous prefer the unseen translations. Mr. Frank Watts[2] among others has endeavoured to construct question-papers of the optional type, calculated to bring out such temperamental differences by the latitude allowed. More recently experiments have been made with pictorial instead of verbal matter: a set of pictures, artistic, humorous, or informative, are placed before the child (picture-postcards supply almost unlimited material for such purposes); and he is asked to arrange them in order of preference or merit; the

[1] One or two American investigators, for example, have attempted to measure ethical discrimination by getting children to arrange a list of offences in order of wickedness. (Fernald, *Amer. J. Insanity*, lxviii., 547; Haines, *Psychol. Rev.*, xxii., 303). Others have attempted to measure moral judgment by noting how often the child singles out moral reasons for certain actions, in preference to reasons of a general or a personal character, (S.C. Liao, *Educ. Contr. Brown Univ.*, III., 1919. *Cf.* Kohs, *J. Delinq.*, VII., i. (1922) "An Ethical Discrimination Test"). Yet another, following up a suggestion of Binet, has elaborated group-tests for honesty and conscientiousness; the children are required to trace mazes with their eyes shut; to fill up and correct completion-tests with the key temptingly handy on the back; to state how much they know of various topics, with the prospect of earning a box of confectionery if they obtain full marks; the measure in these tests is the number of times the child yields to the temptation to cheat or to over-state. (Voelker, "The Functions of Ideals and Attitudes." *Col. Univ. Contrib. Ed.*, 1921; Cady. "The Estimation of Juvenile Incorrigibility," *Journ. Delinq. Mon.*, 1923). In Cady's experiments—the most recent and the most thorough—the results of the tests described correlate with independent estimates of moral character up to .42). Sometimes (as in the last research) the examinee is also given a syllabus of questions relating to his own character: "What kind of amusements do you prefer? Do you get on well with teachers and with other children? Would you like to wear jewellery and fine clothes? What do you think about when you are alone? What would you do if a lot of money were left you?"

[2] *British Journal of Psychology*, XI. 2.

influence of special interests, working quite unconsciously if the pictures have been chosen with care, is nearly always evident; and, if a standard order has been previously obtained, the child's divergences can be measured numerically[1].

48. *The Importance of Observational Methods as Distinguished from Experimental.* Few, however, would as yet pretend that such tests can merit anything more than an experimental interest; and, in their present state, the methods are unsuited for practical employment in this country[2].

In assessing temperament and character, therefore, we are bound to fall back upon the method of observation in place of the method of experiment. The personal interview is one recognised device; and another is the collation of reports submitted by competent observers who have been acquainted with the examinee during a long portion of his life. Both interviewing and reporting has each its own technique; and in either case the technique is susceptible of great improvement by the application of simple scientific principles. Much, indeed, has recently been done by drawing up schedules of facts to be noted and observed[3], and by contriving rating-scales[4] for the registration of such facts in terms of a comparable scheme. More still, no doubt, is to be learnt from the practice of the expert alienist, from the methods and devices employed by the saner psycho-analysts in examining patients who are neurotic, delinquent, or temperamentally deranged. Nowhere is the art of interviewing carried out with such refinement and success

---

[1] An interesting but tentative set of "*Tests of Aesthetic Appreciation*" has been described by Thorndike, in an article with that title, *J. Educ. Psychol.* VII (1916) pp. 509 *et seq.*

[2] See the Committee's comments on these tests in Chapter II, § 64, and Chapter III, § 90.

[3] Of these, perhaps the most suggestive are those given by Webb, "Character and Intelligence," *Brit. J. Psych. Mon.*, 1., and Hoch and Amsden, "Guide to the Descriptive Study of Personality," *Rev. Neur. Psych.*, xi., 577, *Cf. Psych. Rev.*, xxi., 295.

[4] On rating persons either by "relative position" or by reference to "key-subjects" (a method elaborated with some success by the psychologists of the American Army) a rich literature has grown up. See, among other references, *The Personnel System of the U.S. Army*, vols. i and ii; Scott, *Psych. Bull.* xv. (1918): Thorndike *J. Appl. Psych.*, ii and iv. (1918 and 1920); and Rugg, *J. Educ. Psych.* xii. and xiii. (1921 and 1922).

as in the consulting-room of a good psychiatrist. But, when all is said, the problems of temperament and character still constitute one of the most difficult and urgent provinces for future psychological research.

XV.—CONCLUSION.

49. These, then, are the numerous fields of application, into which, first in this direction, then in that, the new psychological methods have progressively spread. And this history, brief and cursory as it is, leaves two things indisputably clear. It shows that the science and practice of mental testing have successfully survived the first critical period of development; but that as yet they have by no means approached perfection, or realised to the full their manifest possibilities. The tests and their uses are all still tentative, still experimental. Many of the expectations raised by the first enthusiasts are yet unfulfilled. Partly from the hasty and crude methods which from time to time have been adopted, partly because the whole problem has proved far more intricate than was at the outset assumed, the need for caution has become increasingly evident. On the other hand, the negative conclusions and the pessimistic verdict, announced by early sceptics towards the close of the nineteenth century, have not been confirmed by the further researches of numerous investigators in the twentieth.

In the whole range of education the art of examining is probably the most difficult; and it was not until recent times that due attention was accorded to its proper principles. Gradually, however, as will have been remarked, the later experiments have become more limited. They have inclined to concentrate chiefly upon certain of the more general capacities, few in number and restricted in kind. The rest have been passed over. Until a few years ago, the majority of the work during the present century was, for the most part, directed to the measurement of one particular ability—intelligence. For the more specific intellectual functions—those underlying memory, attention, and the like—for the important qualities of temperament, emotion, and mental character, and for the native aptitudes for particular occupations (as distinguished from attainments due to practice, training, or experience) the few

relevant researches have yielded as yet hardly a single test-method at once simple and trustworthy. But, with all these necessary reservations, the success and the wide-spread use of intelligence tests remain among the most remarkable achievements of modern experimental psychology.

Finally, as regards general procedure, one striking fact emerges. Starting first of all with technical methods remote from the activities of everyday life, the psychologist has slowly approximated, so at least it must seem to the practical teacher, towards the traditional methods of the school and university—to the method of the oral interview and the method of the written examination. Yet this is no mere reversion ; and the years of labour with tests now discarded have been by no means fruitless. From his laboratory the psychologist has brought with him something of value—a scientific technique capable of being rendered increasingly efficient and exact. His test-questions must now be all carefully standardised beforehand ; and he will rely, in evaluating his results, upon statistically elaborated norms. He thus endeavours to make both his procedure and his deductions as simple in practice, as sound in theory, and as precise in point of scientific method, as the complexity of his task and the limitations of his knowledge will for the time being allow.

CHAPTER II.

GENERAL SUMMARY OF THE AVAILABLE EVIDENCE BEARING ON THE PROBLEMS CONNECTED WITH THE VARIOUS TYPES OF PSYCHOLOGICAL TESTS OF EDUCABLE CAPACITY.

THE VARIOUS TYPES OF PSYCHOLOGICAL TESTS OF EDUCABLE CAPACITY.

50. After careful examination of the evidence, and after consultation with several trained psychologists, regarding the proper definition of the expression " Psychological test of educable capacity," we came to the conclusion that the phrase might, for purposes of convenience be properly interpreted as including the following types of tests :—

(i) Tests of "intelligence," i.e., tests designed to measure that general ability which is held by many psychologists to underlie the various special activities of the mind.

(ii) Standardised scholastic tests based on average performance, i.e., tests of attainments in particular school subjects, such as reading and arithmetic, elaborated by actual experiment and statistical evaluation*.

(iii) Such vocational tests, including tests of manual ability[1] as are adapted for use in schools and educational institutions.

(iv) Tests of mental activities of a specialised kind, e.g., tests of memory, perception, attention, imagery and association.

(v) Certain physical tests which have been suggested as a means of assessing educable capacity.

(vi) Tests of such aspects of temperament and character as bear directly on educable capacity.

---

* See footnote on p. 145.

[1] From the strictly psychological standpoint tests of manual ability should be classified with tests of special mental activities, but as in practice they are closely associated with certain kinds of vocational tests, we have classed them here with such tests.

We desire, however, to point out that in our opinion the terminology employed in any discussion of psychological tests of educable capacity in their present state of development is necessarily provisional, and that the distinctions involved in the above classification are themselves founded on hypotheses, and however convenient for the purposes of analysis, should not be interpreted as if they were in any sense finally valid. In other words, we consider that these distinctions are probably best regarded as first approximations to the truth, and as such are of considerable value for working purposes, but possess only provisional validity.

OBSERVED DISCREPANCIES BETWEEN ABILITY AND ATTAINMENT; THE DESIRABILITY OF RECOURSE TO SOME MEANS OF DISCOVERING INNATE ABILITY APART FROM EXAMINATIONS OF THE ORDINARY TYPE.

51. It is obvious, as was pointed out to us by several of our witnesses, that some simple, uniform and trustworthy device for gauging the educable capacity of children is urgently needed. The method at present so widely employed of appraising the abilities of children chiefly on the results of aptitude displayed in acquiring knowledge of the ordinary school subjects is not altogether satisfactory. It is never quite safe, at any rate in young children, to attempt to assess educable capacity by mere attainment as disclosed in the ordinary written or oral examinations, or even by the application of standardised scholastic tests[1]. We discuss in Chapter III the general question of observable defects in existing arrangements for examinations designed primarily to discover potential educable capacity or general ability and promise rather than attainments; so at the present stage we will merely call attention to the incongruities which have been noted in many instances between the undoubted innate abilities of a child and his actual achievements in examinations.

Meagre attainments in school knowledge may be due to a plurality of causes—to chronic ill-health, or temporary illness involving irregular attendance at school, to unfavourable conditions in the home, including insufficient clothing and inadequate

[1] See § 86.

or unsuitable nutrition, to lack of interest due possibly to failure on the part of teachers to appeal to some emotional element calculated to stimulate enthusiasm for a specific subject or group of subjects, to positive absence of the will to learn, and so forth. A high level of ability does not therefore invariably entail a good standard of school work. It has often been pointed out that many highly gifted men and women were regarded in their school days as dull or incompetent. Such dullness or incompetence may have been generally present, owing to a failure of interest in the pupil, or to an arrested development of ability which was subsequently released. In so far as it is the fault of the school it may not be so much the conventional methods of teaching that are inadequate as the ordinary means of diagnosis, consisting chiefly of written, oral and practical examinations. The school should be criticised not so much for failing to adapt itself to such exceptional personalities as for failing to discover them.

THE APPLICATION OF "INTELLIGENCE" TESTS TO SUBNORMAL AND SUPERNORMAL CHILDREN.

52. It is claimed on behalf of "intelligence" tests, whether individual tests or group tests, that if properly applied and evaluated they afford a more objective, more systematic and more trustworthy means of discovering the existence of inborn intelligence and educable capacity in pupils than the ordinary written and oral examinations. We may point out at this stage, that just as there are three types of ordinary examinations, the written, the oral (*vivâ voce*) and the practical (whether in laboratory work or clinical work), so there are three methods of applying "intelligence" tests and vocational tests. Group tests are set in the form of papers like the ordinary written examinations; individual tests are applied orally in a systematic *vivâ voce* examination; while performance tests and tests of manual ability, or of special vocational aptitude are analogous to the practical examination. The essential feature of such psychological tests seems to be that, unlike ordinary examinations, they only postulate knowledge and experience of a very general kind, and are graded or standardised usually on the basis of the age of the children tested or in terms of mental age.

"Intelligence" tests claim to measure general inborn intellectual ability or "intelligence," which is envisaged as a purely abstract potentiality—an hypothetical quantity postulated and defined, like most other scientific concepts, for the convenience of separate measurement. It is claimed that these tests of intelligence render it possible to predict from quite an early age what will be the probable intellectual level of a child when he is grown up. In fact most of the psychologists who gave evidence before us assured us that within reasonable limits such forecasts could be made, and that the ratio between a child's mental age and his chronological age, which was known as the "mental ratio," or intelligence quotient, appeared to remain tolerably constant during the years of growth. The results obtained from the application of the tests are said to show that beyond the stage of puberty—say 16 years of age—inborn intelligence. i.e. general ability, does not develop to any appreciable extent.

The claim, accordingly is made that it is safe to predict, for example, that a child aged 5 with a mental age of 2 who thus has a mental ratio of $\frac{2}{5}$ (= 40 per cent.), will probably attain a mental age of 4 at the age of 10 and a mental age of 6 at the age of 15, on the ground that inborn intelligence does not appear to develop much after 16, and further, on the ground that such a person would never rise above the six-year level and would probably remain mentally defective for the rest of his life.

The advocates of "intelligence" tests further contend that this form of examination has great advantages over ordinary examinations; in the first place, that when properly prepared, applied and marked, the tests neutralise to a very considerable extent the effects of bad teaching and unfavourable surroundings; in the second place that they furnish a measure not of what the child has learned, but of what he can learn, provided that he possesses the temperament and character requisite to enable him to make proper use of the educational opportunities afforded. In other words, that the tests, unlike ordinary examinations, not merely assess the degree of intelligence, but when applied at an early age detect with a fair measure of accuracy, the degree of educable capacity and hence the extent of probable attainment.

The specific claims made on behalf of mental tests by our various witnesses may be summarised as follows :—

(*a*) That they provided a method of comparing children in respect of their inborn capacity and thus of selecting the best candidates for higher instruction, and sifting out defective and dull children for treatment by special educational methods.

(*b*) That in testing general capacity there must be some material on which to work, so that the results obtained by the application of tests of intelligence necessarily depended to some extent on environment, including schooling, but, inasmuch as in the case of the children compared the tests demanded only a minimum of common environmental conditions they afforded a more even chance to all the candidates, children from good and bad homes and from well-managed and indifferent schools obtaining practically equal chances, which was more than could be claimed for any other form of examination. It was, however, pointed out that there was a small percentage of cases, for example canal boat children, where the necessary minimum of common environment was lacking, so that in such instances the usual tests were not applicable. It was suggested that in order to arouse the intelligence of such children some special set of tests, such as the American Army Beta tests, or performance tests might be employed[1].

(*c*) That tests of intelligence provided a basis of prediction regarding the educable capacity of children before they had actually been taught to any great extent ; in other words, that such tests were prognostic whereas ordinary examinations were mainly diagnostic. They indicated with a tolerable degree of certainty the probable development of the " general intelligence " of the child.

(*d*) That these tests, unlike ordinary examinations, afforded an objective standard of comparison based on the performances of many children, and that the personal equation of the individual examiner was thus eliminated.

---

[1] See *Mental and Scholastic Tests among Retarded Children*, by Mr. Hugh Gordon, H.M.I., Board of Education, Educational Pamphlet No. 44 (1923).

(*e*) That in the application of a test of intelligence an accurate age allowance was made, which could hardly be done in ordinary examinations[1].

(*f*) That group "intelligence" tests took less time to work and could nearly always be marked in less time than ordinary examination papers.

It was stated that the evidence so far afforded for the value of tests as indexes of educable capacity fell mainly under three heads :—

(i) Comparison of the results of the tests with the observations of teachers and others ;

(ii) The following-up after testing of a large number of children in their school careers ;

(iii) The correlation shown by experiment to exist between the results of such tests and those of ordinary standardised scholastic (educational) tests in the case of children who had been afforded normal facilities for education.

WHAT IS INTELLIGENCE ?

53. The purpose and function of tests of intelligence is to determine the general ability and consequently the general educable capacity of the individual pupil. The ordinary connotation of the word "intelligence" differs considerably from the meaning commonly assigned to it by psychologists. It is therefore important at the outset to define so far as possible the significance which is attached to the term by psychologists. Though the word is generally employed both in ordinary use and by psychological writers in the English, French, Italian and German languages, its connotation is vague and elusive. It was originally closely akin to intellect, the sole difference between the two being that intellect, which was defined by some of the Schoolmen as the power to conceive universal ideas, meant the

[1] Several of our witnesses indicated that the considerations stated above seemed to follow from the fact that the technique of mental tests of intelligence apparently consisted in a sampling of the relatively simpler processes which enter into all intellectual functions. The ordinary examinations on the other hand sampled more complicated forms of the simpler processes. The latter processes were partly the product of innate factors, but the particular complications were largely the result of particular training.

faculty, while intelligence denoted its actual exercise[1]. Both words became degraded in popular usage, and were frequently understood as covering other cognitive processes, including memory and even perception. In course of time intellect became the more exact term, being reserved for a power supposed to be peculiar to man, while intelligence was regarded as being shared by man with many of the lower animals. It is not therefore surprising to find that very divergent theoretical definitions of intelligence[2] have been given by the various psychologists who have aimed at measuring it.

The diversity of opinion among psychologists regarding the nature of intelligence or general ability may be illustrated by a recent discussion on the matter by fourteen prominent American specialists on mental testing[3]. Terman, for example, defines intelligence as the power of abstract thinking; Thorndike as the power of good responses from the point of view of truth, Buckingham as the ability to act effectively under given conditions. Ruml pessimistically states that the nature of intelligence can at present hardly be discussed at all owing to the vagueness of the terms involved and our paucity of information about the facts. Indeed, the replies disclose a remarkable variety of opinions. Nor do European psychologists exhibit any greater unanimity.

THE VARIOUS HYPOTHESES REGARDING THE NATURE OF THE FACTORS INVOLVED IN GENERAL ABILITY OR GENERAL INTELLIGENCE.

54. It is possible, however, to arrange the various hypotheses about the nature of "intelligence" under three heads; those that envisage it as the effect of a few highly generalised "faculties" or "functions"; those that regard it merely as a convenient term for the average of innumerable abilities, all highly specific, and those that view it as a single central intellective factor common to all intellectual processes.

---

[1] e.g. St. Thomas Aquinas, cf. Alagona, *S. Thomae Aquinatis Theol. Summae Compendium.* I. 79 (10). "An intellegentia sit distincta potentia ab intellectu. R. Non, sed est actus intellectus."

[2] It was Sir Francis Galton, who first brought the concept prominently into notice in England. cf. His two articles on Hereditary Talent and Character in *Macmillan's Magazine* for 1865.

[3] *The Journal of Educational Psychology* for March, April and May, 1921, Vol. XII, Nos. 3, 4 and 5.

(*a*) The first theory which for purposes of convenience may be termed the "faculty" theory was held by Binet. In his earlier articles Binet speaks of tests for perception, attention, memory, and reasoning—all terms of the traditional school of "faculty psychology." Elsewhere in his writings he approximates to the so-called "two factor" hypothesis, and declares that "in intelligence there is a fundamental faculty . . . This faculty is judgment, otherwise called good sense or practical sense, the faculty of adapting one's self to circumstances . . . Under cover of a test of memory we shall have an appreciation of judgment[1]." In an elaborate article on "The Intelligence of the Feeble Minded," published in 1909[2], he appears to regard general intelligence as a complex mental quality involving at least three faculties:—

(i) The appreciation of a problem and the due direction of the mind towards its execution;

(ii) The capacity for making the necessary adaptations to secure a definite end;

(iii) The power of self-criticism.

The second theory according to which the several traits that go to make up the mind are practically independent is held, with modifications, by Professor E. L. Thorndike[3], who regards intelligence as a multiplicity of innate abilities that are related in varying degrees. He admits that there is a positive relation between desirable single traits in a single individual. "Having a large measure of one good quality increases the probability that one will have more than the average of any other good quality." The fact that a child has pronounced inborn ability in arithmetic indicates that he will have more than the average innate ability in geography, and even that he will be above the average in his moral qualities, but it is not certain that he will be. Thorndike has since suggested that there are three main types of inborn intelligence, viz., intelligence for words

---

[1] *L'Annee Psychologique* (1905, transl. Kite, pp. 42–44).

[2] A. Binet and Th. Simon, L'Intelligence des imbécilles in *L'Année Psychologique* (1909), pp. 1–147, summarised by Terman in "*The Measurement of Intelligence*," p. 45. In a book published in 1909, Binet defined intelligence succinctly as follows: "Comprehension, invention, direction and power of criticism ("*censure*"), intelligence lies in these four words." *Les Idées Modernes sur les Enfants*, p. 118.

[3] *Journal of Educational Psychology*, XII, 126.

and abstract ideas, motor intelligence or skill with the use of the hands, and social intelligence or the ability to get on well with one's fellows. These three types are positively related, but not necessarily in a high degree. A similar belief in the relative independence of intellectual abilities appears to be held in a modified form by Professor Godfrey Thomson[1], who informed us that he did not believe in the existence of a factor called general ability or general intelligence, and considered that the statistical work of those who supported that theory was of doubtful validity.

The third hypothesis which appears to be at present the most widely accepted, in a modified form is that of a central intellective factor. Many of the supporters of this view would probably accept Stern's definition of intelligence as " general adaptability to the new problems and conditions of life[2]." Dr. Cyril Burt, in his evidence, urged that general intelligence was by far the most important of the factors involved in general educational ability. " Intelligence " manifested itself in a number of different ways, but he regarded it more as a single complex quality rather than as a group of independent elements. It was best measured by tasks requiring the voluntary maintenance of attention, quick and accurate learning (in the broader sense of the word, namely, adaptation to relatively novel conditions), and on the higher mental levels, reasoning. These should, perhaps, be regarded rather as modes of general ability than as elements entering into general ability as component factors. Professor Spearman, the first and foremost supporter of this unitary theory, prefers to call the general factor " g," without committing himself to the view that " g " is precisely identical with intelligence as popularly understood[3]. He writes that " possibly this central factor is some general fund of mental energy, which again depends upon a general fund of some sort of brain energy[4]."

---

[1] See Appendix IX, p. 236. cf. Prof. Thomson's statement of his sampling theory of ability in *Essentials of Mental Measurement*, by Brown and Thomson, 1921, pp. 188–192.

[2] *Journal of Educational Psychology*. XII, 127.

[3] *British Journal of Psychology*. V, pp. 51–84.

[4] See also his recent work on *The Nature of Intelligence and the Principles of Cognition*. Macmillan & Co., 1923.

It is obviously impossible for us in this Report to discuss at length the very difficult and technical questions involved in the various theories regarding the nature of intelligence or general ability and its relation to the various specific manifestations of ability. We have, however, quoted in Appendix IX extracts bearing on this problem from the evidence submitted to us by various psychologists.

In spite of the divergence and apparent inconsistency of the various theories regarding the nature of "general intelligence," many of the individual views seem merely to lay stress on one or other aspect of a complex whole. As is pointed out in Chapter I[1] the differences of view are by no means so complete or irreconciliable as they once appeared. Probably in any one intellectual act, factors of at least three orders play an essential part—the "central factor" of "intelligence," the "group factors," which an older psychology would have termed faculties, and "specific factors" entirely limited and independent. The different views seem merely to over emphasise one particular type of factor, and to ignore or under-rate the part played by the other two. Most psychologists would probably agree that "intelligence" is a general mental ability operating in many different ways, given as part of the child's natural endowment, as distinct from knowledge or skill acquired through teaching or experience, and more concerned with analysing and co-ordinating the data of experience than with mere passive reception of them. As, however, this term "intelligence" is employed by psychologists in a technical sense, we have placed it in inverted commas to indicate that it is used with a special meaning.

WHAT DO TESTS OF "INTELLIGENCE" MEASURE?

55. Though, however, we could find no general consensus as to what intelligence is, almost all our witnesses, teachers and psychologists alike, were unanimous in their definitions of what it was not. All are agreed that intelligence does not cover temperament or character; and that, therefore, the important personal qualities of will, feeling, and emotion are not dealt with by tests of intelligence. Secondly, they were agreed that it does not cover *acquired* attainments; hence, tests of intelligence give no indication of what a pupil has learnt in reading,

[1] Chapter I § 15, footnote [2].

spelling, arithmetic, or in any of the higher school subjects. Thirdly, it seems generally agreed that any narrow or limited talent, available for only one type of intellectual work, is not to be named intelligence in this sense. Intelligence is regarded by the majority as a common component entering more or less into all intellectual activities, or, as some would prefer to phase it, as the common level of all particular intellectual performances. What tests of "intelligence" measure, therefore, is inborn, all-round, intellectual ability, using the word "intellectual" in a loose sense to include practical activities as well as theoretical, but to exclude processes of emotion and qualities of character.

These tests aim at measuring this ability through different processes according to the mental level of the child. With the youngest infants little more than simple sense-perception and simple movement can be tested. With the older and brighter children, the commoner tests aim at measuring "intelligence" through the perception of more abstract relations, such as analogy, immediate inference, and the like. But at nearly every level, except perhaps the lowest and the highest of all, the capacity for learning by experience, and (particularly when a time limit is imposed) for learning quickly by a brief experience, is one of the commonest manifestations of "intelligence" to be exercised in these tests.

It should be emphasised that these tests do not pretend to discover all the intellectual qualities that appear to be present in great artists, great musicians, great poets, and masters of literary form. Such specialised forms of genius imply not only the possession of a high level of general intelligence, but also the possession of a specialised talent in an unusually high degree. Without high intelligence no poet could be great. But without the special poetical qualities—imagination, visualisation, intensity of feeling, or whatever they may be—no great man could be a poet. Further, as psychologists as well as teachers have constantly pointed out, there are children of first-rate "intelligence" who cannot express themselves clearly at any length on paper, and who fail to do themselves justice in verbal answers to verbal questions. Such children are likely to do ill with such oral tests as those of Binet or Simon, or with such group-tests as those in the first series[1] used with the American

[1] The "Alpha" tests: the second, or "Beta," series contains tests of a more practical, non-verbal type.

Army and since copied so freely in this country. On the other hand, such children may have very considerable aptitudes in practical or mechanical directions; and with these, if their ability is to reveal itself through any measurable test, some test of a manual or so-called "performance" type must of necessity be used.

A short series of psychological tests cannot, of course, pretend to greater accuracy than the considered judgment of an experienced master or mistress, observing the individual pupils of the class over a period of weeks or terms. They afford only one kind of evidence; and, though even the competent teacher may gain from them something which he could not otherwise obtain—namely, some degree of standardisation or comparability for his judgments—no tests of "intelligence" can completely take the place of observation and insight. They are supplements, not substitutes.

It cannot be too frequently emphasised that even psychologists themselves do not trust solely for their conclusions to the mere quantitative marks from a mechanical set of tests. In the diagnosis of mental deficiency, of backwardness, or of supernormal intelligence, they endeavour always to obtain data and reports dealing with the general conduct and character of the child. There are, it is true, psychological tests which aim at measuring or discovering other qualities besides the intellectual—namely the emotional or quasi-emotional aspects of personality, and the many important traits of character which bear on educable capacity. But, whatever may be thought of the efficiency of such temperamental tests when applied under laboratory conditions, they are in their present state of development unsuited for practical use in school. We are accordingly of opinion that the data afforded by the application of group tests or individual tests of "intelligence," or both, should always be considered in association with the information regarding individual pupils available from other sources, e.g., school records, personal estimates of teachers, achievements in ordinary examinations, medical data, if any, and information regarding home conditions and parentage.

It must always be remembered that character as well as intelligence is, in actual school life, a factor of the highest importance in determining the response of any individual pupil to instruction in any branch of study. Furthermore, there will

probably be some pupils who, though they show little aptitude for the ordinary school studies, have a decided bent, say, in an artistic or mechanical direction. If, therefore, we agree that what "intelligence" tests measure may safely be regarded as a factor operative in all mental operations that can by any latitude be called intelligent, yet at the same time it must be remembered that such a factor is not the sole factor, and that in every distinct kind of mental operation special factors are nearly always involved. In some kinds of mental operation the general factor is of dominant importance, while in others the special factor is more prominent. For operations of the latter kind the intelligence tests alone would have little diagnostic value.

Furthermore, if satisfactory and trustworthy tests of memory, perception, attention, and of such aspects of temperament as affect educable capacity could be devised, it seems certain that the data thus afforded would modify appreciably the data yielded by "intelligence" tests.

THE MAIN PRESUPPOSITIONS UNDERLYING THE USE OF TESTS OF "INTELLIGENCE" AND THE INDISPENSABLE CONDITIONS WHICH MUST BE COMPLIED WITH IN ORDER TO ENSURE THE VALIDITY OF ANY SET OF "INTELLIGENCE" TESTS.

56. There was general agreement that the main presuppositions underlying the use of "intelligence" tests were as follows:—

(i) That there were certain mental factors which remained more or less constant during the lifetime of individual human beings.

(ii) That methods of examination had been discovered or could be discovered by which those factors in any individual could to a great extent be ascertained and differentiated from the results of training and education.

It would seem from the evidence submitted to us that the educable capacity of a child at any period of his or her life may

be assumed to depend on mental factors of two kinds :—(*a*) inborn psychological abilities of a relatively elementary and general nature ; and (*b*) acquired capacities of a more complex and specific character, chiefly memories and habits, such as particular items of knowledge and particular forms of skill. There has been much discussion in recent years as to whether nature (inherited capacity) or nurture (training and environment) is the more important. It is evident, indeed, as Binet frequently pointed out in his articles, that even a child of the greatest potential intelligence can never become highly intelligent in surroundings affording scant opportunity to learn. On the other hand, it is almost certain that a feeble-minded child can never become really intelligent however favourable his environment and however skilled and patient his teachers may be. This applies still more to children with special defects. For example, a deaf mute, whatever may be his inborn ability, will probably grow up apparently feeble-minded unless special methods of instruction are employed to educe and develop his native intelligence. In spite of the hopes that were once entertained regarding results obtainable from special instruction in Schools for the Mentally Defective, a feeble-minded child can seldom, if ever, be converted into a normal child.

It is clear that when a group of individuals are brought up in a similar environment, with almost equal opportunities, the differences disclosed among those individuals by means of intelligence tests are due to differences in native ability. Conversely difference of environment may be so wide as to obscure or even traverse the difference of ability which these tests are intended to elicit. For example, the results of intelligence tests applied to children coming from different environments and living under different conditions, some of whom were underfed and ailing, while others were well nourished and healthy, would necessarily afford untrustworthy evidence of inborn capacity and the value of such results would be *pro tanto* diminished. It follows that the results would be to some extent invalidated, if the persons were drawn from environments that were widely dissimilar or had been subjected to widely different conditions of life.

On a first view it might seem impossible to attempt to measure and appraise the amount or character of an inborn capacity, or

group of capacities that manifest themselves solely through learning, as such capacities can be measured only indirectly through what has been acquired. It should, however, be borne in mind, as was pointed out by several of the most eminent psychologists who gave evidence before us that the distinction between inborn and acquired capacities is mainly a matter of inference. Innate capacity may be indirectly gauged with considerable success by measuring the acquired capacities in a group of children with substantially the same experience. Psychologists infer from differences in acquired intelligence differences in native endowment when they compare individuals in a group of children who have had common experiences and note the differences in the performances and attainments of those individuals.

It follows, therefore, that a valid test of intelligence must be based on elements appealing to the common interests and within the common experiences of the group of persons tested. For example, a careful examination of the Binet tests in their original form shows that the separate tests were arranged on the basis of the common experiences of urban children of varying ages. In no instance was a test employed that was based on peculiar conditions or unusual facilities for learning. Tests for any given age were applied on the assumption that all normal children might have been expected to learn the things with which they were ordinarily in contact. Further, no test of intelligence can properly be regarded as valid unless the individual tested has had reasonable opportunities to learn about the various element. involved in the test and has also been interested in learnings For instance, the American Army "Alpha" tests, which during the last few years have been extensively applied in the educational institutions of the United States, are reported to show in nearly every instance higher average marks for men and boys than for women and girls. The apparent inference that the intelligence of men is on the whole rather superior to that of women is obviously invalid because the tests in question were originally designed for application to soldiers, and refer largely to matters in which women were not likely to be interested.

It was accordingly pointed out that whenever a test was used for diagnostic purposes certain assumptions were made. In the first place, it was assumed that the child had passed his

life in a certain environment. It followed, therefore, that the simpler the environment assumed by the test, that was to say, the more it resembled the environment in which the great mass of the child population lives, the more reliable was the test for application to different age-groups, different social classes, different parts of the country, and so forth.

For example, the Stanford Revision of the Binet Simon Scale assumed that every child tested had come into contact with twentieth-century civilisation and had since the age of not less than seven years spent a certain portion of each year within the walls of an Elementary School class-room. If he had not done so, and to the extent to which he had not done so, the test might yield a result which was not wholly valid. Hence a pupil's prolonged absences from school must to some extent be taken into consideration by the "intelligence" tester.

Secondly, the test assumed that the child had made as much use as his "intelligence" had permitted of the stimuli afforded by such an environment. He might not have done so. His low rating on the "intelligence" scale might then be due in part to an abnormality of temperament.

Our witnesses also emphasised the fact that in order to secure valid results it was most important that the tests should be applied and marked in a rigidly uniform manner. The inconsistent and paradoxical results not infrequently obtained by teachers and others who had applied the Binet tests, or other types of individual tests in a rather unsystematic way, were due to failure or neglect to observe uniformity in application or marking, or in both.

Binet repeatedly calls attention to the importance of careful and uniform administration and marking in the application of his own tests[1], and American psychologists in their writings frequently point out that the significance of results derived from tests may be gravely impaired by lack of uniformity and care in application and scoring. There was general agreement among the psychologists who gave evidence before us that tests, especially individual tests, are more satisfactorily administered if applied by a person specially trained for the work.

---

[1] *L'Année Psychologique*, (1911), pp. 166–168.

THE MERITS AND DISADVANTAGES OF THE BINET-SIMON TESTS.

57. As we have already indicated, the children for whom tests of "intelligence" are especially useful are those who are definitely above or below the average. The celebrated Binet-Simon Scale, as finally revised by Binet himself in 1911, or in its later modifications, especially those known as the Stanford Revision, the Yerkes Point Scale, and Burt's adapted Scale for English children[1], may be regarded as the model of all individual tests of "intelligence" which have been devised up to the present. It has been extensively employed in this country since about 1910 as an aid to the discovery and special treatment of mentally defective and sub-normal children, and to a less degree of super-normal children. It has also been used as an aid for internal classification in Elementary and Special Schools, and to a very much smaller extent in Secondary Schools. The Binet Tests have also been employed for the purpose of checking and supplementing the results obtained by the ordinary examinations for free places and scholarships in Secondary Schools, and for admission to Central Schools. In fact the scale is now widely recognised and used as a convenient mental foot-rule. Bearing this in mind, we took special care to ascertain in some detail the opinions of our witnesses on the merits and defects of the Scale and its variants, and we give below a full summary of their views. There appears to be general agreement that the revised Binet-Simon tests are of real use for the following :—

(1) As an aid to the discovery and classification of sub-normal and mentally defective children.

(2) For detecting super-normal or specially bright children up to the chronological age of 9 (i.e., the mental age of 12), but not beyond.

(3) As tests of progress and retardation in general capacity and for the purposes of classification of scholars on entrance to Elementary Schools or on transference to Senior Departments, or other types of School.

(4) The Binet Scale as adapted by Drummond and others are of use in Schools for children defective in the special senses, e.g., the blind or deaf[2].

---

[1] Burt, *Mental and Scholastic Tests*, pp. 24–68.

[2] W. B. Drummond, *A Binet Scale for the Blind* (reprinted from the Edinburgh Medical Journal). There is also an American form of the Binet scale adapted for deaf children.

While most of our witnesses drew attention to certain defects in the Scale, which are set out below, there was, on the other hand, general agreement as to its merits, which may be stated as follows :—

(1) It was claimed that the Scale was comparatively simple, and that the technique of applying it could be acquired after some training and practice by teachers and school doctors, provided it were explicitly recognised that no final validity could attach to the diagnosis of a layman who had not received a thorough training in psychological method. Some witnesses, indeed, claimed that with a relatively simple scientific invention, such as the Binet tests of intelligence, a layman might accomplish much in the way of rough preliminary diagnosis, which was formerly possible to none but the expert. It was, however, pointed out that the use of such a device by persons who had not had much psychological training must necessarily be somewhat mechanical, and confined and limited in its scope.

(2) The apparatus required for the application of the Scale was simple and inexpensive.

(3) The Binet tests had been much used all over the world, had acquired an extensive vogue in the English speaking countries, and had thus been to a great extent standardised. They might, therefore, justly be regarded as more universal than any of the other individual tests in use.

The defects attributed to the Binet Scale may be summarised as follows :—

(1) It was pointed out that these tests had been originally devised as a scale of measurement to assist the administrative authorities in Paris in examining school children suspected of mental deficiency, and recommended for transfer from the ordinary Elementary Schools to special classes. Indeed Binet, himself, had expressed the opinion that his scale would be chiefly useful for application to mentally defective children[1]. Many of our witnesses, therefore, while admitting that the Scale was of great service as applied to defective children were disposed to doubt whether it was really suitable for general application to

[1] *L'Année Psychologique* (1908), p. 85.

normal and more especially to super-normal children. The Scale was admittedly uneven, and it was more difficult to obtain extra points as the higher end was approached. There was also the disadvantage that the mental processes tested were not the same from year to year. On the whole it was generally agreed that while the Scale was tolerably accurate for the diagnosis of low grades of ability, it was, on the other hand, relatively useless for the diagnosis of high grades of ability among older children. One witness, who had conducted extensive experiments in the use of the Stanford revision of the Scale, pointed out that he found that it ceased to measure accurately the "intelligence" of the brighter children after the age of about 14 years, though it continued to measure the "intelligence" of inferior children with sufficient accuracy.

(2) Those witnesses who had made the most extensive use of the Scale were agreed that it was too linguistic for English children, having been originally designed for French children who received a more elaborate training in the use of language. Tests for later ages involved to a large extent considerable range of verbal imagery. In consequence the somewhat rare type of defective child sometimes described as the sub-normal verbalist, who could talk at considerable length though often irrelevantly, but could do little, possessed a distinct advantage over children who showed little aptitude for the use of written language and whose vocabulary was limited, but who, nevertheless, might possess some mechanical aptitude ; and even more frequently the child who was weak in the comprehension and use of language (if only through some temporary nervousness) might, unless special precautions are taken, be severely handicapped.

(3) Several witnesses drew attention to the fact that a child's proficiency in the Binet tests represented the complex result of numerous intermingling factors. In addition to the two essential items—his inherited "intelligence" and his chronological age—his performance in the tests would be affected by numerous subsidiary conditions, such as industry, goodwill, keenness, emotional stability, information acquired at school, social environment, and sex, which would inevitably impair or improve the result.

(4) Most of our witnesses thought that the most potent of these subsidiary conditions was educational opportunity. Many

of the tests—some of which were withdrawn by Binet in his final revision of 1911[1]—were pure tests of school attainments. For example reading, writing and dictation were learnt in English lessons ; counting, addition and subtraction of money in arithmetic lessons ; drawing from the copy and drawing from memory in drawing lessons ; if, therefore, a child's apparent capacity were assessed solely by the Binet Scale, it must depend in no small degree upon his class in school and his educational opportunities in the past. Conversely, it was urged that, where children's educational opportunities had been normal and equal, a child's school class must depend upon his apparent capacity, as, in theory at all events, the individual pupil was classified on entrance to the Elementary School, and promoted year by year in accordance with what he had already learnt and by what he seemed likely to learn in future.

(5) Several witnesses maintained that the results of the Binet tests were affected by social conditions, whether due to environment or heredity. For example, children from favourable homes were said to prove more responsive to tests than children from less favourable homes.

(6) Several medical witnesses pointed out that mental fatigue was a factor which entered into the test examination in the case of nervous, mentally defective and epileptic children.

(7) A few witnesses had found that children became familiar with the Binet tests, and in certain cases were specially prepared for them. This applied even to mentally defective children.

(8) It was pointed out that a relatively large amount of time was needed in order to apply Binet tests adequately to individual children, as compared with the time required for the collective examination of a whole class in ordinary school subjects.

(9) Several witnesses drew attention to the fact that some of the original Binet tests, having been devised for Parisian children, were not particularly suited for application to rural children in England, as they presupposed a knowledge of facts peculiar to urban life.

(10) Several medical witnesses had found that the language of the Binet Scale required to be adapted to meet the needs of provincial dialects and colloquialisms in various parts of England and Wales.

---

[1] *L'Année Psychologique* (1911), p. 146.

THE OPINIONS OF MEDICAL EXPERTS ON THE VALUE OF THE BINET-SIMON SCALE AND ITS MODIFICATIONS AS AN AID IN THE DIAGNOSIS OF MENTAL DEFICIENCY IN CHILDREN.

58. Our medical witnesses informed us that the Binet-Simon tests had been extensively used as an ancillary method of examining children suspected of mental deficiency since about 1911, more especially in the London area. It was pointed out that these tests were recommended for use by school medical officers in making inquiries into the mental condition of feeble-minded children in the Reports of the Chief Medical Officer of the Board of Education for 1912 and 1913, where it was suggested that the teacher in the ordinary Public Elementary School should select retarded and backward children and in co-operation with the school medical officer should determine by careful examination which children needed subsequent examination by the Binet tests[1]. The actual data obtained by the application of the Binet tests to children suspected of mental deficiency formed only a part of the evidence on which the official diagnosis of cases under the Acts was based[2], physical condition, educational attainments, family and personal history and environment being all taken into account in the consideration of such cases, the aim being to ascertain whether the child was fitted to his normal surroundings in life. At the earlier chronological ages the Scale both in Binet's own final edition of 1911 and in the later editions of other psychologists had been found to be a useful guide, especially when supplemented by standardised scholastic tests in reading and in simple arithmetical processes. Our medical witnesses stated that in view of the inadequacy of the Binet tests for children at later ages it was the practice to utilise other tests in addition, for example, Burt's reasoning tests, association

---

[1] Report for 1912 (*Cd.* 7184), pp. 372 foll., and Report for 1913 (*Cd.* 7730), p. 321 foll. A salutary warning was added against the use of the tests by anyone not fully understanding the proper method, and the conditions and appliances necessary for trustworthy results.

[2] *Education Act*, 1921, section 55, and *Mental Deficiency Act*, 1913, section 2.

tests, and performance tests of various types[1]. On the whole, however, they were of opinion that the Binet Scale had proved of great value as a supplement to the medical and pedagogical methods hitherto employed in the diagnosis of supposed defective children, and that its general adoption for these purposes had as it were stabilised the intelligence tests applied by certifying officers and by teachers in Special Schools.

MERITS AND DEFECTS ATTRIBUTED TO THE AMERICAN MODIFICATIONS OF THE BINET SCALE.

59. *Yerkes' Point Scale.*—Professor Yerkes' Point Scale[2] consists of 20 exercises, of which 19 are taken from the Binet-Simon series ; the remaining one is an analogies test of the type first devised by Dr. Burt. Most of Binet's information tests are omitted. Partial credit in marking is given to the various achievements of the children tested according to their merit, and not invariably by the pass or failure method of Binet. The maximum in the Point Scale is 100, and differences in standards, ages, sex and social status, norms have been calculated in terms of the average number of points scored. These standards, not having the fixity of an age scale, can be readily readjusted and quickly revised if necessary. Dr. Burt and other witnesses regarded this as an undoubted advantage, though they pointed out that the same principle could be applied to the Binet Scale in its usual form by merely counting the number of tests passed, either actually or by implication, and, if necessary, assigning fractions for partial success. By using for partial performances entire points or marks in place of fractions certain specific tests in the Point Scale are assigned a maximum larger than unity and thus carry greater weight than others. Dr. Burt pointed out, that, if this emphasising of certain tests were determined by

[1] See *Annual Report of the Chief Medical Officer of the Board of Education*, 1920 (*Cd.* 1522), p. 100 foll., and *The Health of the School Child*, being the Annual Report of the Chief Medical Officer of the Board of Education, 1922, pp. 111–112. See also the interesting historical account of the use of such tests up to 1910 in Dr. F. C. Shrubsall's Report on methods of testing mental deficiency, printed in *Report of 81st Meeting of the British Association*, 1911, pp. 195–214. *Cf.* also Dr. Shrubsall's article in *School Hygiene* for August, 1921.

[2] Yerkes, Bridges, and Hardwick, *A Point Scale for measuring ability*, 1915.

their diagnostic significance the modification might be of great value, but that for English children the value and maximum mark, as suggested by Yerkes, did not at all correspond.

Brevity is secured by the omission of most of Binet tests that depend, like reading and writing, upon instruction in school, or, like the coin tests, on special experience. On the other hand, several rather inferior tests are still retained, e.g., suggestion, and the comparison of faces, weights and lines. Dr. Burt, who discussed the Point Scale method in some detail, was of opinion that, while of much theoretical interest, this particular revision of the Binet Scale on the whole represented for English children no great advance on the original either for the purpose of examining border-line cases of suspected deficiency at the usual age of entrance to Special Schools, or for testing supernormal and scholarship children at later ages.

The Stanford Revision and Extension (by Professor L. M. Terman)[1].—The outstanding merit of this revision is the inclusion of many well thought-out tests designed for children of higher ages. Other improvements effected in the original scale are the addition of better or more numerous examples of certain types of test ; the allowance of more numerous trials ; and the provision of a more definite method of marking, for example, partial credits for partial success in certain tests.

Binet had expressed the child's intellectual power by giving his mental age in relation to his chronological age. Yerkes, in his Point Scale, indicated the same facts by giving the total points scored by the individual in comparison with the average points scored by normal children of the age of the child tested.

Terman, following Stern and others, uses a somewhat similar method, in stating capacity in terms of the " mental ratio," or in American phraseology, the " intelligence quotient," which is obtained by dividing the child's mental age by his chronological age in order to eliminate the actual chronological age. This method of indication has its advantages, but is also open to certain objections. The chief value claimed for the mental ratio is that it seems to express the child's inborn intelligence in a more or less absolute fashion. It is in fact intended to indicate his actual intellectual capacity irrespective of his age.

---

[1] L. M. Terman, *The Measurement of Intelligence*, Harrap & Co., 1919.

Terman and most other psychologists, as the result of their own experiments, maintain that the intelligence quotient, with possible slight changes, remains constant throughout the life of an individual, at any rate up to the period of puberty.

Some of our witnesses who had made a careful study of the Stanford Revision and its application, seemed to think that on the whole this contention was substantiated by the facts, though it was probable in some instances, at least, that the child's intelligence quotient might vary from year to year, and that at times it might have a tendency to increase and at times to diminish. It was pointed out that while the intelligence quotient served a useful purpose in indicating to the teacher the probable intelligence of an individual pupil at each successive stage of his school career, and was important in helping to forecast the extent and character of his success in school studies, it should never be employed for classifying pupils without also taking into consideration their actual mental and chronological ages. It was obvious that children who had the same intelligence quotient might be far apart in actual acquired knowledge because of differences in mental and chronological ages.

Furthermore, it was a highly abstract and succinct method of expressing the results obtained from tests and should, for practical purposes in schools, be supplemented by notes on the manner in which each individual pupil had "attacked" the tests. It should always be borne in mind that in the present state of development of the tests the value of such "intelligence quotients" was purely relative and in no sense final.

Several witnesses maintained that for purposes of instruction children should be grouped partly on the basis of their acquired attainments and, to a lesser degree, on the basis of their chronological age. It should be added that certain considerations render the Stanford Revision tests less suitable for English than for American children. The principal shortcomings of the Revision from this standpoint may be summarised as follows :—

(1) The older children on whose performances this Revision of the Binet Scale was based by Terman appear to have been of a somewhat higher intellectual level than the average child in ordinary Public Elementary Schools in England.

(2) The proper administration of the revised tests demands almost twice as much time as the application of the original Binet Scale.

(3) The use of American coins and American money values renders it difficult to use the money tests in English schools with their existing age assignments and in their present form, even after the obvious alterations in the names of the coins have been made.

(4) The Scale still retains a strong linguistic bias. In this connection, several witnesses expressed the opinion that more performance tests might advantageously have been substituted for some of the less efficient linguistic tests.

ADVANTAGES AND DEFECTS ATTRIBUTED TO INDIVIDUAL TESTS, OTHER THAN THE BINET SCALE AND ITS MODIFICATIONS.

60. Our witnesses informed us that on the whole the tests devised by Professor De Sanctis, of Rome[1], though occasionally of use as supplementary tests in the diagnosis of mental deficiency, were hardly applicable to normal children, for whom indeed they were not intended, though the general method of the tests was of value for children who could not read or write. Several witnesses were of opinion that the only set of individual tests other than the Binet Scale and its modifications which were suitable for application to English children was Dr. Burt's reasoning tests for children of 7 to 14 years of age. It was stated that as compared with the Binet tests these required more particularly the exercise of the higher intellectual processes. On the other hand, they failed to gauge certain aspects of intelligence and were in that respect inferior to the Binet tests. It seemed accordingly desirable that they should be used in conjunction with the tests of other aspects of intelligence, e.g., tests of a performance type.

[1] See Chapter I, § 30.

THE GENERAL CHARACTER OF GROUP TESTS OF INTELLIGENCE AND THEIR VALUE FOR DETERMINING EDUCABLE CAPACITY.

61. There is an important distinction, in respect of their method of application, between individual tests and group tests. Individual tests are of such a character that they can be administered only to one individual at a time by one investigator. The Binet tests, or tests for a special capacity like pitch discrimination, may serve as types. Group tests are of such a character that they can be administered simultaneously to groups of individuals, larger or smaller according to the circumstances under which the testing is carried out, the size of the group being, however, immaterial as far as the character of the test is concerned. Group tests in this sense[1] are obviously highly desirable for testing on any extensive scale, and therefore raises very important practical problems.

As has been shown in Chapter I the development of group tests is of very recent date[2]. They were in the first instance based on materials of the verbal rather than the performance type, and they are still, on the whole, chiefly linguistic though by no means exclusively so. Naturally the group tests applied to very young children are largely of the performance type such as the Porteus Maze. The Alpha tests[3] employed in the American Army, which were published after the end of the War, have been very extensively used in the United States for students in Universities, Colleges and High Schools. The first of these is a Directions test to determine ability to execute commands. The second is an arithmetical problem; the third consists in selecting from three possibilities the best reason for a statement; the fourth represents a list of words associated in pairs, and the examinee has to determine whether these words are associated by the principle of opposition or of likeness. The fifth is a disarranged sentence, and the examinee is required to put the words in their proper order so as to make sense. The sixth is a number completion test

[1] The expression "group test" has occasionally been used by some writers in the sense of a test for "group" abilities as opposed to "special" abilities. It is never employed, however, in this sense in the present Report.

[2] Chapter I, § 31.

[3] See Appendix VIII for examples.

in which a series of numbers has to be continued according to the number indicated in the part of the series given. The seventh deals with analogies, or mixed relations; the eighth with range of information. One of the most important tests for revealing individual differences is the completion test devised by Ebbinghaus in 1905 for the purpose of investigating the fatigue of a school day in Breslau[1]. The original test consisted of a paragraph in which words with syllables omitted were presented to the examinee, who was required to fill in the omissions. Terman has a high opinion of this type of test, which, he says, discloses fundamental differences in the thought processes.

The Analogies or Mixed Relations test first used by Dr. Cyril Burt in 1911 consists in presenting three words in a series, the first and second of which bear a certain relationship. The examinee is required to supply a fourth word that bears the same relationship to the third word as the second does to the first. It is claimed on behalf of this test that it is suited for discovering some of the more complex forms of intelligence[2].

This test is typical of a large number classified under the general name of Association. Controlled Association tests include, beside the Analogies, associations of part with whole or conversely; the genus with the species or the reverse; a word with its opposite and so forth. This test admits a variation by the substitution of pictures or designs for words. The substitution test which determines quickness and accuracy of learning by substituting for one set of characters another according to a key is also included in group "intelligence" tests.

Among group tests which are extensively used may be noted those which deal with Vocabulary, and which are really included under Range of Information, and those which elicit response to verbal orders. The latter have now been modified so that they can be used with pencil and paper.

Another type of group test commonly employed at present is an exercise in the simple processes of arithmetic. This type involves concentrated attention, mental alertness and, in some

---

[1] The theory underlying this "combination" test was expounded by Ebbinghaus in an article published in 1897. *Zeitchr. f. Psych, und Psys. d. Sinnesorg.* xiii., p. 401.

[2] See examples in Appendix VIII.

instances, a considerable amount of reasoning power. The marks obtained usually exhibit an appreciable degree of relationship to general intelligence Obviously, however, it must depend in part at least upon instruction and practice in arithmetic, and probably also upon a special arithmetical ability which is partly independent of general intelligence.

It has also been claimed that some of the reading tests, particularly those constructed by Professor E. L. Thorndike, are able to measure some of the higher mental abilities. The examinees are required to read a paragraph, and then answer certain questions concerning it with the paragraph still before them.

It will be seen from this brief description of the character of some of the more important elements in the group tests at present in use what degree of intelligence is required in order to answer them[1]. Several of our witnesses were of opinion that, on the whole, the more complex factors of judgment, inference and of logical analysis were not extensively involved.

It would appear that most of the existing group tests are dependent on verbal material and that very few of them have as yet been adequately standardised. Several witnesses, however, thought that it would not be difficult to devise and standardise tests with the available non-linguistic material and yet call into play the same higher mental functions as the verbal tests. The Porteus Maze and the Healy Picture Completion were probably the best efforts in that direction. It was repeatedly pointed out that group tests, though they did not afford the same insight into the child's mind as did individual tests, were the only type that could in practice be applied when large numbers of children had to be examined.

As we note in the next section, very few group tests have as yet been devised which are suitable for application to children under 10 years of age. In practice, therefore, the only types of test which are suited to young children under 10 are individual tests of intelligence and standardised scholastic tests in simple subjects such as reading and elementary arithmetic applied orally to each individual pupil.

---

[1] One of the best known sets of group tests devised by English psychologists is the Northumberland tests, constructed by Professor Godfrey Thomson. See Appendix VIII.

On the whole, our witnesses were of opinion that the data afforded by the use of group tests, though of very considerable value, gave only general evidence, and that in all doubtful and borderline cases they should be supplemented by individual tests, which usually produced more accurate results.

THE RELATIVE MERITS AND DISADVANTAGES OF GROUP TESTS AND INDIVIDUAL TESTS OF INTELLIGENCE.

62. It was generally agreed that one of the chief advantages of group or collective tests, which were set in the form of written papers, was economy of time, while the principal merit of individual tests, which were applied at an oral interview, was the insight afforded into the child's mind by personal observation of his answers and general attitude. Moreover, group tests were on the whole more finely graded than individual, and if they had been properly elaborated and standardised beforehand their application demanded relatively little special training, whereas for the satisfactory application and nterpretation of individual tests, a training in experimental psychology, in the technique of application, and in the use of statistical methods, was indispensable.

It was pointed out that the proper application of individual tests imposed a considerable nervous strain on the examiner, and that they were frequently of a type which might admit of special preparation beforehand, whereas in group tests the human factor, as represented by the examiner, was more likely to be kept constant. Up to the present, very few group tests had been devised which were suitable for application to children under 10 years of age[1]. In any case they were best suited for older and brighter children, and were not well adapted for very young children who had but recently acquired the art of writing. Conversely, the individual tests hitherto constructed were described as being especially useful for young children under about 10 years of age. Those devised up to the present were, however, of comparatively little use for older children, particularly above the age of 15 or 16, while, on the other hand, group tests had been constructed which were suitable for children of 16 and upwards, and even for adults.

[1] Cf. Burt, *Mental and Scholastic Tests*, p. 221, "With children under the age of 10 and below Standard IV the results (of group tests) will correlate less highly with intelligence."

Again, group tests implied to a considerable extent the use of pen or pencil in writing, drawing, underlining, etc. When speed of performance was used as a measure of efficiency an error in rating might creep in, for all pupils did not handle pen or pencil with the same ease and facility. It followed that the results obtained were generally of the nature of statistics, and in order that the group records might be of *diagnostic* value, it was necessary in most instances to test pupils individually. In other words, the data obtained from group tests were probably best regarded as first aids towards mental diagnosis. We should mention, however, that one distinguished psychologist assured us that he had found the results of written group tests to be almost as efficient as those obtained from individual oral tests above the level of Standard II in Elementary Schools, though it was necessary to call attention to the dangers of expecting the same norms from group tests as from individual tests. The general opinion of our witnesses was that, for most practical purposes, group tests should be applied first, and should be followed up, when the services of persons competent to administer individual tests were available, by the individual testing of children whose performances in the group papers had been noticeably abnormal, whether below or above the line. Several witnesses directed attention to the fact that none of the group tests at present in use served to gauge with sufficient precision the following aspects of intelligence :—(*a*) ability to control and concentrate attention ; (*b*) ingenuity and inventiveness ; (*c*) practical judgment ; (*d*) ability to manipulate mental imagery.

The great advantage of the individual tests was that they furnished the expert with concentrated material for observation.

It was pointed out that it was of the highest importance that an examination should call forth maximum effort on the part of every pupil. A group or collective test (in this respect resembling the ordinary school examination) sometimes failed to do this. However carefully the examiner arranged his question it was inevitable that a few pupils should fall short of their best, and the most accurate instructions would sometimes be misunderstood. In individual testing, on the other hand, while the examiner rigidly adhered to standardised procedure he yet found it possible to vary the " attunement " of the pupil in such a way that he secured real effort and trustworthy results.

THE PLACE OF THE INTERVIEW IN THE APPLICATION OF INDIVIDUAL TESTS AND ITS TECHNIQUE.

63. Several witnesses thought that the interview, so far as it related to children of school age, was, as at present conducted, frequently rather a test of social opportunity and upbringing than of intellectual capacity. The technique of interviewing had in the past been too empirical and had not been sufficiently studied for its own sake. Indeed, there was little doubt that the psychology of the interview was relatively an unexplored field. In regard to the possibility of teaching the technique of interviewing in connection with psychological tests, several witnesses thought that it should be possible to obtain help from the principles of modern psychotherapy. In general it was urged that more attention should be devoted to the place of the interview both in ordinary examinations and in the application of psychological tests with a view to drawing up a clear statement of the data which could and should be obtained by its use. One witness went so far as to say that the interview as ordinarily conducted was likely to lead to injustice because of the fluctuations in the examiner's powers as well as in those of the candidates. He himself to some extent had found it possible to render judgments arrived at in the result of an interview more objective by methods similar to those employed in the American Army Rating Scale, combined with information given to those conducting such interviews in regard to the fluctuations which were to be expected by reason of "sampling," or chance.

In general, our witnesses were of opinion that the interview would always have its uses, as the most satisfactory results could only be obtained by combining written group tests with oral individual tests. It was pointed out that group tests were at present of little value for ascertaining qualities of temperament and character, but that the expert might in an interview glean useful information, though for a really safe estimate it would be sounder to rely on the judgment of experienced observers who had been in contact with the child during longer periods. One witness thought that it would be good policy to aim in paper examinations at measuring intellectual ability and in interviews at appraising the force and quality of the interests and pursuits of individual candidates. The interview might be employed with advantage in further scrutiny of doubtful and borderline cases which had already been picked out by means

of the written answers. Another witness suggested that manual tests might in some instances be introduced into the interview, as they could seldom be applied simultaneously to a large number of children. It should be mentioned that Binet attached great importance to the technique of the interview in the application of his individual tests, and gave detailed suggestions regarding the general conditions of the oral examination. Dr. Cyril Burt thought that the technique of the interview was susceptible of great improvement by the application of simple scientific principles. Something had already been done to this end by drawing up questionnaires of facts to be noted and observed and by devising rating scales for the registration of such facts in such a way as to make them comparable.

HOW FAR DO TESTS OF INTELLIGENCE THROW LIGHT ON CHARACTER OR TEMPERAMENT, AND HOW FAR ARE THEY AFFECTED BY THEM?

64. Our witnesses have admitted that tests of "intelligence," while measuring in some degree intellectual ability, are also dependent on the opportunity which the child has to learn and his interest in learning. There are several other considerations involved in the ability to perform them, the principal of which is perseverance and concentration or, in other words, the capacity to hold the mind down to a task and keep the attention alert and concentrated. The tests were not, indeed, primarily designed to reveal character, as Binet pointed out in the following passage of an article written in 1908 :—" Our examination of intelligence cannot take account of all those qualities, attention, will, regularity, perseverance, teachableness and courage, which play so important a part in school work, and also in after life ; for life is not so much a conflict of intelligences as a struggle between characters, and we must in fact expect that those children whom we consider to be the most intelligent will not always be those who are the most advanced in their studies[1]." On the other hand, the exercise of will power, implying a certain measure of concentration and persistence, was involved in the performance of an intelligence test, so that, if applied by a skilful and observant tester, tests of intelligence might incidentally throw light on certain aspects of character, but of course only to a slight and limited degree.

[1] *L'Année Psychologique* (1908), p. 77.

But, as Binet pointed out in several of his articles, the tests may yield unsatisfactory results owing to peculiarities in the character or temperament of those to whom they are applied. For example, it is a difficult matter to apply them satisfactorily to highly nervous children, or to clever children who are disposed to suspect that there is some catch underlying the apparently simple questions put to them by the tester.

Various attempts were made by Binet to devise as adjuncts temperamental tests for the purpose of assessing conscientiousness, suggestibility, and accuracy of reproduction[1]. In the last few years systematic efforts have been made, more especially by American psychologists, to test the feelings and the will. Pressey has endeavoured to detect repulsion and forms of fear by asking the subject to select from a prearranged list of words those that have for him a special meaning, or suggest special dislike or irritation. It is claimed the Porteus Maze affords to some extent a means of measuring recklessness and impulsiveness and, as Dr. Burt has pointed out, variability in repeated tests of almost any simple type as indicated, for example, by the standard deviation, appears to be partially correlated with instability. On the whole, however, the psychologists who gave evidence assured us that no tests of temperament could at present claim to have passed beyond the stage of tentative experiment.

Several witnesses of wide teaching experience, assured us that it was improbable that gifts of character, such as earnestness of effort and the will to succeed, would ever in school work triumph over the insuperable obstacle of an extremely low intelligence quotient. Clever children would learn if they wanted; the intellectual attainment of dull children, however much they tried, would never develop beyond a certain point. On the other hand brilliant children frequently had grave faults of character which, if not corrected, limited effectively their educable capacity. Thus, the development of satisfactory quantitative tests of temperamental characteristics would provide a subsidiary estimate which would have to be taken into account in ambiguous and borderline cases.

Some of our witnesses suggested that teachers should observe more systematically weakness of character as well as weakness of intellect. Often what seemed an intellectual failing was really due to some temperamental defect.

---

[1] See § 46. *cf.* also A. Binet *L'Etude Experimentale de l'Intelligence* Paris, 1903.

It was urged, also, that there was a definite need for teachers to be able to judge the more elusive characteristics of character and temperament. In the playground, as well as in the classroom, they had excellent opportunities for such observation; and many of them, owing to their special gifts and special experience, were quick and penetrating judges of youthful character. But they sometimes seemed to find difficulty in setting down what they had noticed in clear and accurate terms. Clearly formulated upon sound psychological principles, such judgments, it was asserted, would be of much greater value than the data afforded by any existing tests of temperament or morality.

CONNECTION BETWEEN EMOTION AND GENERAL EDUCABLE CAPACITY.

65. In a memorandum on the question of tests of educable capacity in relation to the emotions, Mrs. S. S. Isaacs, basing her views upon recent psychological research, held that the emotional factors in mental development must be taken into account on the ground that the level of scholastic or practical achievement might be affected by temporary or permanent emotional conditions. She pointed out that the same test might have very different effects on different children according to the "complexes" which are aroused in the unconscious self. She held that certain mental tests of general capacity should be revised and that it was desirable that those who applied such tests should receive some training in the observation of emotional reactions, since such training yielded a fuller understanding of the emotional life and afforded a means of investigating those causes of disability and failure which had their origin in unconscious conflict.

These views have been placed before several of our witnesses, both psychologists and teachers, and it has been pointed out:—

(1) That if emotional disturbances or aberrations are liable to affect the results of mental tests they are equally liable to affect the results of the ordinary scholastic tests. It is quite true that "inhibition" may deleteriously affect the answer to an individual group test. For example, the use of a particular word in a vocabulary test may prevent the child from answering that particular question. It must, however, be remembered that the psychological examiner employs

several tests, so that inhibitions of this type are unlikely to persist over the whole paper of group tests.

(2) That if emotional disturbances are liable to modify the results of mental tests they equally disturb and modify the whole course of school life, and that therefore the results obtained by the application of mental tests, while they may not indicate the maximum capacity of the emotional pupil, are likely to afford a trustworthy and practical indication of the educational potentialities of the individual.

(3) That as a matter of fact, those familiar with the views of the Freudian School who have been keen to observe evidence of the existence of "inhibiting complexes" have been impressed by the rarity of the existence of cases of obvious and grave emotional disturbances of Freudian type in younger school children. They admit that with the onset of puberty these disturbances are liable to be present and so may very possibly affect the results of mental tests just as they affect the results of ordinary scholastic examinations, both written and oral.

THE VIEWS OF OUR WITNESSES ON THE VALUE OF STANDARDISED SCHOLASTIC TESTS BASED ON AGE PERFORMANCE FOR GAUGING EDUCABLE CAPACITY.

66. Several of our witnesses were disposed to believe in the principle underlying standardised[1] scholastic tests and pointed out the desirability of drawing up a standard series of such tests and obtaining norms[1] of performance therewith at successive ages more especially for children over 11.

Dr. Cyril Burt thought that it would be a much simpler task to draw up a standard series of scholastic tests and to obtain norms of performance with them at every age than to construct tests of "intelligence." He considered that the value of such scales would be very great provided their purpose and significance were rightly understood. There was general agreement regarding the importance of obtaining for such tests not only a norm of average performance, but also some measure of the degree of individual variation.

[1] See Appendix IV (Note on Standardisation and Norms).

Dr. Drever was of opinion that a trustworthy series of tests of attainment in the various school subjects was eminently desirable, for while it was true that the more essential and subtle effects of education did not lend themselves to direct testing, it was none the less certain that there were important results which did so lend themselves and that those results furnished a good, though not an infallible guide, towards an estimate of the success of education as a whole. In some instances there was not much difficulty in devising satisfactory tests, as, for example, in the fundamental processes of arithmetic: in other subjects, such as in reading or composition, the task of constructing satisfactory tests was harder, and scarcely any of those hitherto devised could be regarded as more than tentative. Some work done by a Research Committee of the Scottish Educational Institute indicated the possibility of satisfactory testing with respect to composition, but much research was still necessary. He thought that Dr. Burt's work on the subject of scholastic tests[1] had furnished tests and norms which must supersede previous work for this country. Caution, however, must be exercised in using Burt's norms which were designed for London children. This was particularly observable in certain subjects, such as spelling for example, where other circumstances might exert a noticeable influence on the ability of the children. Thus one observer had found in an Edinburgh school an advance of a year, or a year and a half on Burt's norms for spelling, due probably to the fact that the spelling problems were easier for the Edinburgh children, owing to the more phonetic character of Scottish pronunciation.

Professor Percy Nunn regarded favourably the efforts which were being made to establish norms of performance for pupils of different ages in the examinable parts of the school curriculum, but he strongly deprecated any proposal to impose their use upon the schools by external authority. If used voluntarily by a teacher as a means of finding out how his pupils stood in a given subject, they might often have great value, but their compulsory use would lead to undesirable narrowing of the curriculum in poorer schools, to unhealthy pressure and to the cramping of initiative. He also pointed out that, if standard norms of performance were established it would be necessary to vary them from time to time as the curriculum and teaching methods changed.

[1] Burt, *Mental and Scholastic Tests.* pp. 257 foll.

Other witnesses were of opinion that standardised scholastic tests based on age performance would serve the same function in the realm of attainment as was served by the Binet tests in the realm of intelligence, namely, the ordering of children of a given age on the basis of comparison with the average child of that age. Furthermore a comparison of the results derived from the application of the Binet tests and their modifications and the scholastic tests respectively in the case of individual children would afford material from which interesting conclusions might be drawn.

We incline to the opinion that such tests even if brought to a high stage of development would not in themselves be sufficient to discover the existence of innate capacity as distinct from acquired knowledge, but would require to be supplemented by tests of intelligence.

Furthermore it seems probable that pupils could be prepared for standardised educational tests with greater ease than for "intelligence" tests, so that it would be very necessary to have numerous sets of standardised tests in each general school subject.

THE VALUE OF VOCATIONAL TESTS (INCLUDING TESTS OF MANUAL ABILITY) IN DETERMINING EDUCABLE CAPACITY.

67. It was pointed out that the vocational end was one of the earliest purposes with reference to which the desirability of some adequate system of testing was recognised, and that historically, therefore, vocational testing should be regarded as prior both to mental and educational testing, though scarcely any real progress had been made with vocational testing before the development of individual or differential psychology towards the end of the nineteenth century. Vocational tests might be roughly classified in accordance with the degree of congruity between the actual problem set to the examinee and the subsequent situations in view:—

(*a*) Vocational tests frequently had reference to situations of a definitely limited range, e.g., those involved in printing, or in driving an electric train. In such cases fitness for a particular vocation might be determined by presenting the candidates with precisely or approximately the same type of situation which the vocation in question required and estimating

quantitatively their capacity for dealing with such a situation; subject to the condition that it must not presuppose technical skill and knowledge which had yet to be acquired. It was evident that on the whole tests of this type were more suited for application in the workshop, factory, office, or in the laboratory, than in an educational institution.

(*b*) In a second type of vocational or quasi-vocational tests the aim was to discover not so much how the examinee could deal with a specific kind of situation, as whether he possessed specific capacities. In such cases fitness for a particular vocation might be determined by testing the candidate for those special capacities which were requisite or desirable in that vocation.

Some witnesses, after pointing out the limited possibility of vocational tests designed to anticipate the type of situation which the vocation presented, recognised a growing tendency to rely on the second method of testing candidates for those special qualities which were, or were supposed to be, requisite or desirable in the particular vocation. Other witnesses, however, emphasised the greater value of the first class of tests. Incidentally, several witnesses expressed the opinion that the most pressing problem of vocational testing at the present time was an adequate psychological analysis of the various requirements of the industries under consideration. The psychologist could probably select or devise suitable tests, if only he had a precise knowledge of the processes involved in a particular vocational task and of the special capacities on which ability in that vocation depended. Unfortunately, however, such knowledge was in many instances lacking. When psychological tests had been made more satisfactory, they could be used to discover persons having a relatively high degree of the capacities required for a given occupation. Such persons were more educable in the directions indicated by the tests in the sense that they could attain a higher degree of proficiency after training than those less highly endowed.

One witness thought it was desirable to devise particular tests for each trade that was at all specialised. It might be possible to have the same tests of general "intelligence" for all, but after that various groups of trades would require tests for special capacities. It would be necessary to apply a series of such special vocational tests in order to ascertain the particular occupation for which a child was best fitted, but it

was important, in the first place, to discover the general level of "intelligence." The opinion was expressed that, whereas in occupations which were not mainly of a routine nature the more "intelligence" a person possessed the greater the probability of success; on the other hand, a positive lack of general "intelligence" might be an advantage for some kinds of mechanical work. Data in support of this conclusion have actually been published[1].

Vocational tests are at present being employed in this country—

(*a*) in Chocolate Factories,

(*b*) in Engineering Works,

(*c*) in a Dressmaking Firm,

(*d*) in a Polytechnic and in several Junior Technical Schools.

(*e*) in connection with the placement of children at school leaving conferences.

Several witnesses pointed out how great would be the value of trustworthy tests of vocational tendencies for entrants to Trade Schools and to Central Schools. If, for example, such tests could be used to assign pupils to the industrial or commercial type of Central School respectively they would be of great service. We were informed that vocational guidance by the aid of vocational tests was already being attempted in certain Elementary and Trade Schools in the London area[2].

THE NATURE AND VALUE OF PHYSICAL TESTS AS SUPPLEMENTARY TO THE PRECEDING.

68. The connotation of the expression "physical tests" is somewhat ambiguous, but our Sub-Committee defined it as covering the ordinary medical examination of the individual together with the measurement of certain bodily reactions and traits which might possibly be closely connected with and therefore might throw light upon mental activities.

---

[1] Cf. Otis A.S. "The Selection of Mill Workers by Mental Tests." *J. of Applied Psychology*, 1920. IV, 339–341.

[2] See §§ 41 and 42.

It is generally agreed that the state of health affects capacity for mental work, and, therefore, that tests of general health and vitality do measure something connected with the intelligence, that any obvious departure from normal health in a candidate should be duly noted, and that the medical reports upon candidates should be available for reference in determining the action to be taken upon the results of tests of intelligence.

In this connection physical tests include the following :

(*a*) The measurement of certain physical dimensions such as the length and breadth of the cranium, the blood pressure, lung capacity, etc.

(*b*) Measurement of the acuity of sense perception—of eyesight, hearing and touch.

(*c*) Measurement of quickness of response to a stimulus (reaction time), muscular power, as measured by the dynamometer ; muscular resistance to fatigue, as measured by the ergograph ; accuracy of aim and general dexterity, i.e., measurements indicating the grade of co-ordination between the higher nervous centres and the muscular system.

Of our witnesses Dr. William Brown considered that these physical tests had some slight value ; there was, according to his experience, a certain correlation between the results of physical and mental tests, although admittedly the correlation was low.

The general opinion of our witnesses and of psychologists as a body is that physical tests are of little service in estimating educable capacity but are possibly of higher value in determining certain orders of vocational capacity. The aesthesiometer has been extensively used to measure the sensitory discrimination of the skin. The test most frequently applied to detect muscular control is the tapping machine, but the results up to the present have proved unsatisfactory. McDougall's Plunger apparatus is sometimes employed to measure accuracy of aim as well as speed of performance. One witness who employed both these pieces of apparatus found that neither of them afforded a trustworthy measure of either practical ability or general ability.

In Dr. Cyril Burt's opinion these tests in general had been primarily designed to measure physical qualities, and as a consequence were of subordinate value in the determination of educable capacity. It might be possible and useful to establish a series of tests specially devised to determine both general intelligence and physical capacity with particular reference to determining the aptitude of pupils for definite vocations and for special forms of education in preparation for these vocations.

The only carefully thought out series of tests known to us along these lines are the interesting researches of Dr. Mumford, of Manchester, upon the boys of the Manchester Grammar School, and on a limited number of University students. These researches are still in progress. So far as they have gone they indicate a definite correlation between breathing capacity and place in school during the earlier ages and in the lower forms, where boys with a wider scope of breathing, and so, presumably, with a more rapid absorption of oxygen and discharge of carbon dioxide, were, in general, found taking the higher places in class. This relationship between respiratory capacity and mental ability became, however, less noticeable in later years, from sixteen upwards.

Another series of observations directed to correlate the marks awarded to some 250 candidates at the School Certificate Examination with their physical growth and breathing capacity has been made and is in process of evaluation by Dr. Caradoc Jones, Lecturer in Mathematics at the University of Manchester, which, when completed, will be presented before the Medical Research Committee which has aided in these researches.

The application of breathing tests to a number of second-year medical students appeared to show a similar correlationship between higher breathing capacity and high standing in University work. These observations of Dr. Mumford[1] need to be

[1] Mumford, Alfred A. *The Relation between Mental and Physical Efficiency of Boys at the Manchester Grammar School.* Trans. Manchester Statistical Soc., 1921 (Dec. 14), pp. 23–45. *Estimation of Physique and Stamina for School Purposes,* Lancet, 1915, I, p. 115, and *Estimation of Physical Fitness in Terms of Respiratory Movements of the Several Regions of the Chest.* Lancet, 1922, I, p. 478. Other papers on the same subject have been published by Dr. Mumford, in conjunction with Mr. Mathew Young, in *Biometrika,* 1923, and in the *Journal of Scientific Physical Training,* Vol. 15, No. 43, p. 12 and its extension, the *Journal of School Hygiene and Physical Education,* Vol. 15, No. 44, p. 45.

repeated on a larger scale with various orders of boys and girls from Public Schools, and Secondary and Elementary Schools situated both in towns and in the country before any generalisation can safely be made.

THE DANGER OF SPECIAL PREPARATION OR "COACHING" FOR THE TESTS.

69. Several witnesses admitted that they had found in practice that children sometimes became familiar with individual tests, and in some instances had prepared for them beforehand. Such previous preparation might be the result either of communication from one child to another, or of special "coaching" beforehand by teachers or other persons. There was little to be feared in regard to the possibility of communication from one child to another. When individual tests were applied, the pupil would generally be given, in the space of about half-an-hour 30 to 40 different tests, most of which comprised 3 or 4 parts. Consequently his after recollection of those tests would probably be very vague, and in practice it was found that children could give only very confused accounts of such tests afterwards.

As regards group tests of "intelligence," it was claimed that there was little danger of special preparation beforehand, since there was ample scope for indefinite variation, both in the form and in the actual matter of such tests.

It was further pointed out that at present teachers seemed, as a rule, to take the keenest interest in the tests, and it was consequently improbable that any very serious attempt would be made to spoil them by "coaching." It could not, however, be assumed that this state of affairs would continue indefinitely, but in any case it was considered probable that coaching for group tests would defeat its own end, in view of the large number and variety of tests of that type.

One witness had found in administering the tests that, when children had been practised beforehand, the correlation was much more trustworthy and that a higher average was obtained. Several other witnesses recommended that every child should receive a little preliminary exercise, when either individual tests or group tests of "intelligence" were applied, in order to put them at their ease.

Another witness suggested that all candidates might be put on a level in the matter of familiarity with tests, by being given a preliminary test just before the actual examination.

On the whole, our witnesses were of opinion that the dangers of coaching were not very great, and that a little preparation beforehand might even have good results. It was pointed out that a child who had received special preparation beforehand on one test could not do another well unless he possessed the capacity tested, as routine procedure was impossible in such group tests as "analogies," "completions," and "directions." In any case, a much more trustworthy result could be obtained by applying the tests several times than by applying them once only. If this procedure were adopted, the existence of special preparation beforehand could probably be detected, and the tester would be in a position to observe the actual effects of practice and appraise the results accordingly.

It was further pointed out that in testing for general ability the investigator was frequently more concerned with the character of the mental process than with the result, and that when teachers realised this they would recognise the relative futility of any special preparation beforehand. At the same time, most of our witnesses were agreed that, if in the future individual tests were brought into more general use, it would be of great importance to construct several sets of properly standardised tests as well as to provide alternative tests in the same scale.

As regards standardised scholastic tests in specific school subjects, such as reading and arithmetic, it was pointed out that teachers in Elementary Schools were hardly likely to attempt to prepare pupils, entering at 7 or 8 years of age, beforehand for such tests. On the other hand, it was admitted that, if in the future standardised scholastic tests were devised for children in the top forms of Elementary Schools and in Secondary Schools, it would be necessary to construct a considerable number of standardised tests for the several ages in order to reduce the danger of coaching.

On the whole, our witnesses seemed to be of opinion that it was more probable that older children could be successfully coached for standardised scholastic tests than for tests of "intelligence."

SUGGESTIONS BY WITNESSES IN REGARD TO THE TRAINING REQUIRED IN ORDER TO ENABLE TEACHERS AND OTHERS TO APPLY AND MARK TESTS OF "INTELLIGENCE" AND STANDARDISED SCHOLASTIC TESTS.

70. We adopt the suggestions made to us by several witnesses that the work in connection with psychological tests of educable capacity may conveniently be classified under four heads:—

(*a*) The construction of the tests;
(*b*) The application or administration of such tests;
(*c*) The marking of the results;
(*d*) The interpretation thereof.

The psychologists who gave evidence before us were agreed that the devising and standardising of all types of psychological tests should be entrusted only to trained psychologists. It was, however, pointed out that it was most desirable that psychologists should keep closely in touch with school work and school conditions when constructing such tests. In regard to individual tests of intelligence there was general agreement that for the satisfactory application of such tests and for the accurate interpretation of the data obtained by their use, careful training in experimental psychology, in the technique of applying the tests and in the use of statistical methods was indispensable. It was repeatedly pointed out that results obtained from the application of individual tests by untrained persons should always be received with the utmost caution and were in fact as a rule almost devoid of scientific value. Some of our witnesses thought that experienced teachers, provided they possessed the necessary gifts of personality, could be trained sufficiently in a short time, say two months, to apply the Binet-Simon tests, whether in their original form or in one of the later revisions.

On the other hand, there was general agreement that the correct interpretation of the results obtained by the use of individual tests could only be satisfactorily carried out by persons who had received scientific training in a psychological laboratory for at least two years.

As regards group tests of intelligence, our witnesses were of opinion that their satisfactory application, when they had once been properly elaborated and standardised, demanded relatively little special instruction, though it was important

that the instructions given to the person who actually conducted them should be carried out with precision. The technique involved, however, was relatively simple and many witnesses thought that a teacher should be able to acquire the necessary knowledge by attending a short course in educational psychology with special reference to the significance of group tests and the technique of administering and marking them. On the other hand it was pointed out that the interpretation of the results of group tests was quite as difficult as the interpretation of the data obtained by the use of individual tests, and that it should only be undertaken by persons who had had the necessary special training in a psychological laboratory.

In regard to standardised scholastic tests our witnesses were of opinion that the technique of applying them was not difficult and could as a rule be acquired by teachers after a little instruction and practice.

As regards vocational tests there was general agreement that they should only be administered and interpreted by trained specialists.

RECOMMENDATIONS BY WITNESSES ON THE DESIRABILITY OF ESTABLISHING A CENTRAL ORGANISATION TO TRY AND DIRECT NEW TESTS AND TO COLLATE EXPERIENCE.

71. There was general agreement among the psychological experts who gave evidence before us in regard to the desirability of establishing a Central Committee, or Clearing-House, which could arrange for continuous investigation in various fields over a number of years. Furthermore, now that the first broad outlines of the tests had been worked out, the process of further refinement and standardization could only be effected by the combined efforts of a large body of co-operating investigators, including specialists in every relevant subject, and of independent workers engaged in collecting data upon a scale sufficiently extensive for statistical analysis. Several witnesses suggested that to this end the Board of Education might set up a central committee comprising administrative officers and inspectors (both from the Board and from Local Education Authorities), teachers, school doctors, psychologists with a knowledge of school children and trained statisticians.

The work of a Central Committee, organised on these lines, would be not unlike that of the National Physical Laboratory in its own sphere. The Central Committee could keep in close touch with bodies such as the Education Section of the British Psychological Society, and the National Institute of Industrial Psychology, and the Psychological Departments of the various English and Welsh Universities. In this connection special attention was drawn to the important work, in trying and directing new tests and in collating experience, which is being carried on by the Education Section of the British Psychological Society, and by the National Institute of Industrial Psychology. Dr. Drever further suggested an additional plan of State and local intervention, namely that mental testing should be undertaken as supplementary to the medical inspection of school children, and that every educational area should have its own psychological clinic, but that on the practical side these clinics should be in close connection with the schools, with employment bureaux, appointments committees, and the like, while on the research side they should be in touch with the psychological departments of Universities, and from the vocational standpoint with the National Institute of Industrial Psychology.

CHAPTER III.

THE VARIOUS POSSIBLE APPLICATIONS OF PSYCHOLOGICAL TESTS OF EDUCABLE CAPACITY IN THE PUBLIC SYSTEM OF EDUCATION.

PART I.

THE POSSIBLE USE OF "INTELLIGENCE" TESTS IN CONNECTION WITH ORDINARY EXAMINATIONS INTENDED PRIMARILY TO TEST ABILITY.

72. In Part I of the present chapter we discuss the practical application of psychological tests in connection with various types of examination, in particular free place examinations. Before, however, we begin our examination of that problem we desire to call attention to one general conclusion which has impressed itself upon us with steadily increasing force as our inquiries have proceeded, and which we wish to state at the outset, since it forms an over-riding condition to which all our particular conclusions and proposals must be read as subject. It is that any system of selection whatever, whether by means of psychological tests, or by means of examination, which determines at the age of 11 the educational future of children is, and must be, gravely unreliable. We agree, of course, with the generally accepted view that transference to the Secondary School should normally take place at the age of 11. What we desire to emphasise is the waste of capacity inevitably involved in any arrangement under which a limited number of free places is fixed in advance, and only such children are admitted to Secondary Schools as are successful in winning them as the result of competitive examination, whether of the conventional kind, or supplemented by some form of psychological test.

It is, unfortunately, the case that as long as the accommodation of Secondary Schools and the provision of free places is so inadequate, as in most areas it is at the present time, selection by one method or another is inevitable. In such circumstances the free place examination, which was intended to be a qualifying examination, becomes a competitive examination, and often a competitive examination of an extremely stringent and

exacting kind, since a large proportion of the candidates cannot possibly, whatever their ability, be successful in winning a free place. The results are that a large number of children who are well qualified to profit by secondary education are excluded from it, and that the whole subsequent career of many children is determined by their success or failure in the single competitive examination held at the age of 11. We are convinced that if selection is to be intelligent and its results trustworthy, it should come at as late a stage as possible in the educational career of children, and that any system of selection for higher education at so early an age as 11 is unsatisfactory.

We feel bound, therefore, to point out that no mere improvement in the methods by which selection takes place—eminently desirable though such improvement is—can act as substitute for the large increase in the Secondary School accommodation and in the supply of free places which is so urgently needed.

73. *"Intelligence" tests as a species of ordinary written and oral examinations, designed primarily to discover abilities.* From one aspect tests of "intelligence"[1] may be regarded as a new type of examination, designed primarily to discover educable capacity. The basal idea involved in the term "examination" is substantially the same as that underlying the new tests, which may indeed be envisaged as a form of examination intended to discover ability. The various means employed at present to assess educable capacity fall into two main classes :— (*a*) Tests of attainment in the various school subjects and of ability as revealed in the presentation of such acquired knowledge, and (*b*) tests of "intelligence." Tests designed to discover ability through the medium of acquired knowledge in the various school subjects consist of the ordinary written and oral examinations as at present conducted and of standardised educational tests[2] based on the average abilities which have been determined

[1] "Test" meant in old English the cupel (testum) used in assaying gold and silver alloys or ore.

The fundamental idea in "test," which is that of assaying metals (and later, after about 1800, chemical substances) is much the same as that of "examination," or weighing. Cf. Latimer, *Sermons* (Parker Society), II. 104 (1552). "Calamities be but examinations and proofs." In this connection it is worth pointing out that "examination," derives from "examen," the tongue of a balance. *Scholiast on Persius*, 1. 6.

[2] See §§ 33 to 36 and 66

for the various ages or standards in certain fundamental branches of study. Tests, regarded as a form of examination, differ in two respects from ordinary written and oral examinations designed primarily to discover ability as revealed in the manipulation of acquired knowledge: first they claim to be exactly standardised, and secondly the acquired knowledge which they postulate is of a very restricted character, inasmuch as they aim mainly at detecting inborn ability or mother wit.

The word "examination" is ordinarily used to denote a systematic test of knowledge, or of either special or general capacity or fitness, whether carried out under the authority of some public body, or applied in the form of an internal examination set to the pupils or students of some one school or educational institution[1].

If we classify ordinary examinations in respect of their aim or end, we may delimit them for purposes of convenience into (i) knowledge (or memory) tests, or more exactly tests of the power of restating or reproducing facts and arguments of a kind that may be learned by heart: (ii) tests of capacity or, in other words, of the power of doing something, for example, writing a report on a specific subject with a particular object in view; solving a mathematical problem or making a précis of a written document, writing an essay on some subject within the candidate's range of knowledge, or a test of ingenuity in applying knowledge of scientific principles to some concrete problem: (iii) tests both of ability and of attainments. In point of fact a well conducted examination should test both actual attainment and ability. Nevertheless, for working purposes, there ought to be a fairly clear line of distinction in aim between an examination such as the free place examination, intended to select those candidates who are likely to profit most by instruction in a Secondary School "on the ground of their higher capacity and promise[2]," and an examination such as the First School

[1] Cf. Brinsley, *Ludus Literarius* (1612), V, 48, "That every yeare there be a solemne examination by the Governors of the Schoole." Cf. also *ibid*, XXVIII, 282, "Which worke of continuell examination is a notable quickener and nourisher of all good learning."

[2] See Appendix to the Regulations for Secondary Schools (England), 1922, and Appendix to the Regulations for Secondary Schools, Wales (including Monmouthshire), 1922.

Examination designed primarily to ascertain whether a candidate has reached a certain minimum standard of attainment in three main groups of school subjects.

74. *The First and Second School Examinations.* For our present purpose, which is to consider the possible use of "intelligence" tests as adjuncts to ordinary examinations we are concerned only with examinations designed primarily to discover general educable capacity, through the medium of a certain minimum of acquired knowledge in a few specific subjects.

The First School Examination is defined in the Board's Circular of June 15th, 1917[1], as being one intended for pupils about the age of 16 and designed to test the results of the course of general education before the pupil begins such a degree of specialisation as is suitable for advanced work in Secondary Schools. The condition of passing the examination is that candidates shall have reached the required standard in certain main groups of school subjects. It is obvious, therefore, that a pupil, however great his inborn ability, cannot pass such an examination unless he possesses a certain amount of detailed knowledge in several branches of study. For the same reason, the Second School Examination, which assumes that the candidate, after the stage marked by the first examination, has followed a more specialised course in certain accepted groups or combinations of subjects[1], must be regarded as being largely an examination in a special branch of acquired knowledge. It is true, as we have already noted, that any well conducted examination will test ability as revealed in the use and disposition of acquired knowledge. But as both the First and the Second School Examinations are principally concerned with the testing of acquired knowledge, we do not consider that "intelligence" tests in their present state of development could usefully be employed in them.

75. *Examinations for free places in Secondary Schools, for entrance to Secondary Schools and for admission to Central Schools.* By far the most important public examination intended to discover capacity and promise in young children, rather than to test their attainments, is the examination for free places in Secondary Schools (and in some cases for admission to Central Schools), conducted by the various local education authorities

[1] Circular of June 15th, 1917, Sections 8 (*c*) and 9, and Article 48 of the Regulations for Secondary Schools (England), 1922, and of the Regulations for Secondary Schools, Wales (including Monmouthshire), 1922.

in England and Wales[1]. The aim and character of this examination is defined generally in the Appendix to the Regulations for Secondary Schools, which, after laying down that the minimum standard of qualification for admission to Secondary Schools must be the same for fee paying pupils and for free place pupils, provides that "where there are more qualified candidates than places, a preference may be given, as regards candidates for free places, to the children of those parents who are most in need of assistance in sending their children to a Secondary School, or such other preference as may be specially approved by the Board. Subject to any such preference, priority shall be given both for free and other places to those candidates . . . . who are considered, on the ground of their higher capacity and promise, likely to profit most by instruction in the school[2]." The qualifying entrance test for the admission of candidates who are 10 or over on the first day of the school year in which they desire admission must be suitable to their age and designed to ascertain their fitness to profit by instruction in the school. Pupils admitted under the age of 10 must also pass the required entrance test by the end of the school year in which they become 11. The Board lay down that the entrance test may properly be oral as well as written, and that regard may be paid to reports upon the candidates from the heads of the schools which they have previously attended. Candidates under 12 on the first day of the school year in which they desire admission must not be required to answer written questions except in English and arithmetic. It is further provided that the governing body of the school shall be responsible for the conduct of the entrance tests, except in so far as they may arrange for a local education authority to conduct them. The Board indicate that it is desirable that persons with experience of teaching in Public Elementary Schools as well as those with experience in Secondary Schools shall take part in conducting them, and that an opportunity should be afforded to the head master or head mistress of the school which the pupil desires to enter of expressing an opinion upon the pupil's fitness before he is finally admitted.

---

[1] The total number of free places held in Secondary Schools on the Grant List in England and Wales on 1st October, 1922, was 129,505. It should be pointed out, however, that a certain number of these free places were not awarded by Local Education Authorities.

[2] Regulations for Secondary Schools (England), 1922, Appendix Rule 1 (*d*). Ditto for Wales, 1922, Appendix, Rule 1 (*d*).

The examination for free places in Secondary Schools always comprises papers in English and arithmetic, and sometimes also a general paper. In the areas of some of the larger local education authorities a preliminary qualifying, or "weeding out," examination in English and arithmetic is held in the local Elementary Schools, and the children selected by means of this test are then called upon to sit for a more elaborate examination in English and arithmetic, which is generally held at convenient centres throughout the area.

76. *Expedients adopted by various local education authorities to improve the examination for free places as a test of ability.* It is generally admitted that it is far more difficult to discover ability[1] than to test acquired knowledge and it is not therefore surprising to find that much attention has recently been devoted by many education authorities to the problem of improving the technique of free place examinations. Many local authorities have recently been devoting much care and attention to the conduct of their free place examinations with a view to rendering them a more exact means of selecting those children who, on the ground of their higher capacity and promise, are considered likely to profit most by instruction in a Secondary School. The more important means[2] adopted up to the present to this end may

---

[1] As early as 1877 Mr. Henry Latham pointed out how difficult it was to test ability: "Further, the gauging of ability is a much more delicate matter than the weighing of knowledge, and it can only be entrusted to an Examiner of special skill. He must be able to recognise the qualities which are disclosed by the performances of a candidate, and he should follow the workings of his mind as if he were part of himself. Such Examiners cannot be readily found, though there are plenty who can judge of acquirements; this puts a practical difficulty in the way of selection of persons on a large scale on the score of ability. It increases this difficulty, if the public insist on having a list of numerical marks as a guarantee of fairness." H. Latham, *On the Action of Examinations Considered as a Means of Selection.* Cambridge, 1877, pp. 218-219.

[2] A full discussion of the general character and methods of conducting free place examinations up to 1920 is contained in the *Report of the Departmental Committee on Scholarships and Free Places*, 1920. (Cmd. 968). Several of the arrangements described above have, however, either come into use, or have been considerably developed since the publication of that Report.

be summarised as follows :—

(*a*) Much care and attention has been devoted by many authorities to the setting and marking of the obligatory papers in English and arithmetic in order to render them suitable and reasonable tests for children of about 11 years of age.

(*b*) Some authorities, in addition to the obligatory papers in English and arithmetic, set a general paper which not infrequently contains questions designed to discover native wit and ability as distinct from acquired knowledge—questions which are in fact unstandardised tests of " intelligence."

(*c*) A few authorities in addition to the oral examination also have recourse to an oral reading test, in which marks are assigned for such qualities as accuracy, apprehension of meaning, correct emphasis and expression, and dramatic feeling.

(*d*) Some authorities arrange an oral examination for all candidates, while others hold an oral examination for borderline and doubtful candidates only.

(*e*) Some authorities have taken steps to ensure that marks are subsequently standardised on a systematic plan in order to correct, so far as possible, any peculiarities of individual examiners in the matter of marking.

(*f*) The supervision of the examination is not infrequently entrusted to an external examiner, on whom devolves the task of organising the whole examination, setting the papers both for the preliminary and for the final examination, presiding at the oral examination, and standardising and correlating the marks.

In addition to any or all of these expedients, several authorities have also been experimenting in the application of " intelligence" tests, which are usually set in the form of a separate paper of group tests for all candidates[1] at the second stage in the free place examination. Some authorities also employ standardised individual tests of " intelligence " as part of the oral examination, more especially for doubtful and borderline candidates.

---

[1] In some instances a few group tests have been included in the ordinary papers in English and arithmetic, e.g. in the free place examination held by the West Riding Education Committee in 1922.

It will thus be seen that up to the present the use of "intelligence" tests in connection with examinations for free places has only been one of several means employed to render such examinations more effective.

77. *Evidence regarding certain shortcomings of examinations for free places as at present conducted and the possible use of "intelligence" tests as adjuncts thereto.* It does not come within the purview of our present Reference to make any exhaustive and far-reaching criticisms of free place examinations and examinations for entrance to Secondary Schools as at present conducted[1]. We think, nevertheless, that it may be useful to summarise the opinions of our witnesses on certain defects of the existing system, as their criticisms may help to indicate the directions in which "intelligence" tests might usefully be employed to supplement and check the data obtained from the papers in English and arithmetic which form, as it were, the core of the existing examination. We should, however, say by way of preface to our summary of these criticisms that, so far as we can ascertain, the present examination, having regard to the vast number of candidates who are presented every year from Elementary Schools, works well on the whole[2]. The most serious objection that can be urged against the present system appears to be that some candidates who fail to obtain free places are possessed of greater natural ability than some of the successful candidates.

Several witnesses were of opinion that the ordinary free place examinations, oral and written, admitted some pupils to Secondary Schools who were not suited for that type of education and on the other hand rejected a number of suitable candidates, particularly those who developed late, or who came from remote rural schools, or from small urban schools where owing to difficulties of staffing pupils did not enjoy educational opportunities such as were afforded in many of the larger schools. For example, the head mistress of a County Secondary School stated that she had come to the conclusion that candidates at the head

---

[1] See, however, § 72.

[2] For example, it appears from statistics supplied to us by the London County Council Education Committee in respect of 2,617 pupils admitted to 51 Departments of Central Schools on the result of the Junior County Scholarship Examination, that on the whole the grading in the Central Schools at the end of the fourth or fifth year corresponded with the grading at entrance.

of the free place list were not always those who did best during school life, and that those who were lowest in the list, or who, having failed to pass the examination, entered as fee-paying pupils frequently did better in school than many pupils who entered with high marks. Her general conclusion was that while, with few exceptions, most free place pupils justified the award, the present system of selection did not necessarily discover the best candidates from the Elementary Schools.

The head master of another Secondary School said that it was not uncommon to find that after a year at school not only did the free place pupils reverse the position which they had held at entrance, but that several pupils who, having failed to obtain free places, had entered the school as fee-payers, had considerably outstripped pupils who had won free places. This evidence was corroborated by another head master, who stated that in his school out of 25 free place pupils selected by competition from over 300 candidates every year, it almost invariably happened that two, three or four had to be moved down from the upper to the lower or middle forms, while, on the other hand, several fee paying pupils who had failed to win free places were promoted at the end of their first or second term at school to take the place of the free place scholars who had been moved down. Various considerations were adduced to explain the failure of some pupils to justify their success in the free place examination. The reason most commonly assigned was that the special preparation which, despite all means taken to discourage it, was undoubtedly given in many Elementary Schools[1], enabled some children to pass the examination with distinction. Another reason given was that some children had a talent for "window dressing." The failure of good candidates to pass the examination with credit was sometimes also due to ill-health or nervousness. One head master ascribed the occasional retardation of some free place pupils in Secondary Schools to unfavourable home environment, including not infrequently insufficient nutrition and unsuitable clothing; to lack of ambition; to the absence of the stimulus and special preparation to which some

[1] Cf. Sir Graham Balfour's Report on the County Scholarship Examination for 1920–21 in *Minutes of Staffordshire Education Committee, July 23rd*, 1921, p. 97. "What we seek to avoid is allowing children, for payment or otherwise, to receive a special preparation to enable them to outwit the examiners and to outstrip their less fortunate rivals not by ability but by mark catching."

of the pupils had previously been accustomed in their Elementary School; and to differences in the curricula and methods of instruction employed in Elementary and Secondary Schools respectively.

It was further pointed out that able pupils from remote rural schools or from small urban schools, who owing to difficulties of staffing had not received any special preparation for the free place examination such as was frequently given to pupils in larger schools, were handicapped by the fact that so much weight was in practice assigned to the paper in arithmetic, a subject for which relatively mediocre children could often be successfully "coached" up to a certain point. In practice, the arithmetic paper was often the deciding factor, even though the same maximum of 100 marks were assigned to it as to the papers in English and general knowledge. This was chiefly due to the fact that the divergence between the highest and lowest marks obtained by candidates in the arithmetic paper was often far wider than in papers in English and general knowledge. It had, however, been observed in several free place examinations that where questions in mental arithmetic were not placed at the beginning, but interspersed in the body of the paper, the divergence of marks was not so great. This was doubtless due in part to the fact that some able children, who had had little practice in doing such problems, having entered the examination room in a state of considerable nervous tension, and found questions at the beginning of the arithmetic paper of a type unfamiliar to them, had been unable to do themselves justice. It is true that similar instances of mistaken ways of setting papers both in respect of the nature and the form of the question, and in regard to the position assigned to it in the examination paper, might be cited from papers actually set in English and general knowledge, but in these cases the evil is very much less. We are of opinion, therefore, that quite apart from the question of the value of "intelligence" tests in themselves, it is of very considerable importance that examiners who set papers for free place examinations should endeavour, as no doubt is already often done, to place themselves in the position of the examinees, most of whom have been presented for a public examination at the age of 11 for the first time in their lives, and many of whom have been previously told by their parents, guardians, teachers, and, in effect, by the State, that their whole future career depends on the winning of a free place. We are of opinion that, if greater

pains were taken to frame the written papers on sound psychological lines, and to bring a knowledge of educational psychology to bear on the conduct of the oral examination, the argument for the general use of "intelligence" tests as adjuncts to the free place examination would be *pro tanto* weakened, though such tests would still be of considerable service in borderline cases.

None of our witnesses suggested the substitution of "intelligence" tests for the present free place examination. Most of them, however, were disposed to favour the employment of such tests as ancillary to the existing examination, though one head master expressed the view that the existing examination was as adequate as anything humanly fallible could be, and that the so-called "intelligence" tests used as an adjunct thereto brought to light a precocity which did not stand the test of time.

Several witnesses suggested that the most satisfactory arrangement would be to interpolate group "intelligence" tests into the ordinary papers in English and arithmetic. Others proposed that the scope of the examination should be enlarged so as to include not only written papers in English and arithmetic, but also a paper in group tests of "intelligence," supplemented by reference to the school record, and the head teacher's estimate. It was further suggested that in all doubtful cases individual tests of "intelligence" should be applied.

On the whole, none of our witnesses were prepared to accept the results of any existing "intelligence" tests as by themselves conclusive, but most of them were disposed to think that group tests, if properly administered and accurately marked, would serve to supplement and check the results obtained by the ordinary examination, and would afford useful guidance in doubtful cases. It was generally agreed that in all really doubtful cases the data afforded by the ordinary examination and by group tests should be supplemented by an oral interview at which individual tests might be applied by a trained expert.

78. *Suggestions for the experimental use of "intelligence" tests in connection with examinations for free places.* After consideration of the available accounts of experiments conducted up to the present time by local education authorities in the use of "intelligence" tests in connection with free place examinations, we are of opinion that in areas where the number of candidates

is large, tests of " intelligence " if carefully applied and marked, would be of considerable use as adjuncts to the ordinary papers in English and arithmetic. So far as we can see, the tests might be used experimentally in connection with the free place examination at any of the following stages :—

(*a*) *In the Elementary School before the free place examination.*—A paper of group tests might be set to pupils in Elementary Schools at the age of about 11, shortly before they were presented for the free place examination. The information afforded by the application of these group tests, provided they were carefully marked and evaluated, would be of considerable service, if considered in association with the school record and the teacher's estimate of each individual pupil. It seems to us that the data afforded by group tests applied in the Elementary Schools which the children ordinarily attended would probably be more trustworthy than the results of such tests when they were worked by the children under new and unfamiliar conditions at the examination for free places. Group tests applied at this stage in the familiar surroundings of the local Elementary School might very possibly afford information that might lead to the presentation of some children for the free place examination proper who, on the results of their work in English and arithmetic in the qualifying or " weeding out " examination might have been adjudged unfit to sit for the free place examination.

(*b*) *The use of group tests in free place examinations.* It seems to us that there may also be a field for experiment in the use of group tests of " intelligence " in free place examinations, provided always that the results obtained by the use of such tests be regarded as merely " indicative," or affording a presumption of the presence or absence of native ability. In other words, we do not think that a candidate who had done poorly in the papers in English and arithmetic and had yet acquitted himself creditably in a paper of group tests should be awarded a free place unless he had successfully passed a fairly searching oral examination designed to discover whether the apparent ability revealed by the group tests was really present. In this connection we would point out that several examiners who had applied both group tests and individual tests in connection with free place examinations were of opinion that what was revealed by these tests was often mere presence of mind, or " sharpness " in seeing the point, while in the absence of trustworthy additional

tests of certain important aspects of character which affected educable capacity, valuable qualities such as power of concentration and intellectual sincerity might be ignored.

(*c*) *The use of individual tests at the oral interview held in connection with the free place examination.* We are of opinion that experiments might also be usefully made in the application of individual tests as supplementary to the oral examination, but we think that these should be set only in those cases where a qualified person is available, and that the oral examination should consist mainly of questions of the ordinary type designed to discover mental alertness and certain aspects of character essential to success in a Secondary School, such as the power of continued attention and concentration.

79. *The importance of the oral interview in examinations primarily intended to discover ability.* In connection with the use of individual tests of intelligence which must necessarily be applied at an oral interview, we desire to emphasise our opinion that more attention should be devoted by psychologists and teachers to a clear definition of the precise aims, functions and limitations of oral examinations generally and to the proper development of their technique. We are of opinion that in all examinations which set out primarily to discover ability great weight should be attached to the results of an oral examination.

It is well known that in France special attention has been devoted for many years past to elaborating the technique of those oral interviews which form an integral and most important part of the State examinations for the *baccalauréat*, the *licence*, the *agrégation* and the *doctorat*[1]. It is therefore particularly significant that individual tests of intelligence were first developed by Binet, who built up his Scale by adapting and improving an already existing form of oral interview devised by doctors and teachers with a view to segregating for instruction by special methods children suspected of mental deficiency[2]. Indeed, from one aspect, the Binet Scale may be regarded as a standardised form of oral examination for children between the ages of 3 and 14. It is interesting to note that

---

[1] *Cf.* Board of Education, *Special Reports on Educational Subjects*, Vol. 24 (1911), pp. 287–289.

Girard, *Questions d'Enseignement Secondaire*, Paris, 1905, p. 65.

[2] See Chapter I, §§ 20, 21; *cf. L'Année Psychologique* (1905), pp. 163–336.

attention has recently been directed in several official reports[1] to the importance of the oral interview for older students as a means of discovering not merely knowledge, but intelligence, mental alertness, and real mastery of such acquired knowledge as the candidate may possess.

80. We consider that an oral interview is of special importance in free place examinations and in examinations for entrance to Secondary Schools and for admission to Central Schools, since, as is well known, many children—even those trained under favourable conditions in the best Elementary Schools—cannot always do themselves justice on paper.

Several of our witnesses who possessed extensive experience of free place examinations pointed out that a child of real ability might often fail to do himself justice in a written examination. It was not difficult to explain such failure in children of

[1] *Cf. Report, dated 20th June, 1917, of the Committee appointed by H.M. Treasury to consider and report upon the scheme of examination for Class I of the Civil Service.* (Cd. 8657). Section 29 :—

"The Royal Commission expressed a cautious inclination towards a *vivâ voce* examination, but made no definite recommendation. The Consultative Committee in their interim report on Scholarships, dated 19th May, 1916, say that there should be a *vivâ voce* examination. On this point, as on almost every point of our Report, we are unanimous. We believe that qualities may be shown in a *vivâ voce* examination which cannot be tested by a written examination. We consider that the *vivâ voce* can be made a test of the candidate's alertness, intelligence, and intellectual outlook, and as such is better than any other. The *vivâ voce* examination has been proved by experience to redress in certain cases the results of the written examination. The examination should, of course, be skilfully conducted by carefully selected examiners accustomed to handle young men and to put them at their ease. We consider that the *vivâ voce* examination should not be in matters of academic study, but in matters of general interest, on which every young man should have something to say."

See also Section 36 of the same Report : "The *vivâ voce* should be a test, by means of questions and conversations on matters of general interest, of the candidate's alertness, intelligence, and intellectual outlook, his personal qualities of mind and mental equipment."

Cf. also *Report of Modern Languages' Committee*, 1918, p. 209 : "Oral Examination should be used wherever possible ; and in School Examinations and in Scholarship Examinations it should always be possible."

See also Recommendation 54 of the same Report : "Oral examinations should be more general and should not be used merely as a means to test candidates in speech."

only 11 years of age. Lack of practice in the literary expression of ideas and indeed in the mere art of rapid penmanship, the new and unfamiliar environment of a public examination, native diffidence, nervous tension, and anxiety about the result, on which so much appeared to depend, were common disturbing factors. The play of such factors might be counteracted by the ardour or the happy unconsciousness of the candidate, but even so, some subconscious "inhibition" might be present in his mind which prevented him from doing himself justice in a written paper. On the other hand, a skilful oral examiner with an adequate knowledge of child nature could often put even a nervous[1] examinee at his ease after a few questions, and induce him to talk. Even if the oral examiner inadvertently put to the child an otherwise simple question in a form which was unfamiliar he could readily elucidate it orally in a way which was impossible with an unsuitable or badly set question in a written paper.

We understand that of late years a number of local education authorities have devoted such attention to the conduct of the oral examination as part of their free place examination, and that special care is exercised to put children at their ease and to ask questions designed to detect fundamental traits of intellect and character and to induce the children to talk about their

---

[1] It should be pointed out that due weight must be assigned to other considerations regarding nervousness in youthful examinees. For example, a child may be nervous because he or she possesses a very fine or acute temperament or mental susceptibility. On the other hand, we think that even as applied to children of 11, there is some force *mutatis mutandis* in the arguments adduced in the following passage from the *Report of the Treasury Committee on the Class I Examination for the Civil Service.* (Cd. 8657.) Discussing the *vivâ voce* examination for young men of about 23 years of age, the Committee write (section 29): "It is sometimes urged that a candidate—otherwise well qualified—may be prevented by nervousness from doing himself justice *vivâ voce.* We are not sure that such lack of nervous control is not in itself a serious defect, nor that the presence of mind and nervous equipoise which enables a candidate to marshal all his resources in such conditions is not a valuable quality. Further, there are undoubtedly some candidates who can never do themselves justice in written examinations, just as there are others who under the excitement of written competition do better than on ordinary occasions. We do not consider that it is desirable to forego the *vivâ voce* test for the advantage of a few weak vessels."

special interests in order to afford opportunities for the revelation of any latent promise of ability. One examiner, for instance, who had had great experience in arranging and supervising free place examinations, told us that, though up to the present time he had never employed individual tests of intelligence at the oral interview, his general aim was (1) to discover character; (2) to appraise intelligence, and (3) to give credit for knowledge of facts. He had found the oral examination of the greatest service for discovering able children, who had not done themselves justice in the written examination. In his opinion the oral examination was of far greater value than the written because young children of 11 often did not know how to deal with written papers.

It would be absurd to claim more than partial validity for any system of testing (whether by written and oral examinations of the ordinary type, or by group or individual tests) of intelligence which aims at determining so early as the age of 11 the suitability of a child to proceed to places of higher education[1]. Nevertheless, of all the means at present employed in order to discover latent ability and promise among candidates for free places[2], we are disposed to attach special importance to a well conducted oral examination. Until recently the ordinary oral examination was too often casual and haphazard. The examiner did not as a rule prepare his questions beforehand, or standardise his material, or put the same questions to all candidates; and thus had no ground for the comparison of candidate with candidate. On the other hand the oral interview at which standardised individual tests of intelligence are applied by a trained tester is the most searching and incisive of all forms of examination by means of "intelligence" tests[3].

---

[1] See § 72.

[2] In areas where there are great number of candidates we think that oral examination might be dispensed with in the case of candidates at the top or bottom of the list. Cf. Sir B. Gott's *Report on Examinations for admission to Secondary Schools in Middlesex*, 1923, p. 8.

[3] As has been pointed out in § 52, tests of intelligence in respect of method of application fall into three classes—individual tests which are applied orally at an interview; group tests which are set in the form of written papers; and performance tests, which in the method of their application may be compared with "practical" examinations, though in point of fact many performance tests have been reduced to pencil and paper, and may be included in a written paper like linguistic group tests.

PART II.

THE POSSIBLE USE OF "INTELLIGENCE" TESTS WITHIN SCHOOLS OF DIFFERENT TYPES.

81. Having indicated the possible use of tests of intelligence as adjuncts to certain public, or quasi-public examinations intended primarily to discover inborn ability, we now pass on to consider the possible use of such tests within schools of various types. Here there is a wide and fruitful field for experiment, and it would obviously be impossible in the present state of development of "intelligence" tests to give any authoritative and complete account of the ways in which they might be employed by teachers as a basis of classification and promotion of their pupils, and as a subsidiary means of detecting sub-normal and super-normal children. Relatively few careful and scientific experiments have been made in the application of such tests to pupils in our schools[1], and there is, accordingly, at present no large body of experience to draw upon, such as is available in the United States, where the tests have been extensively employed for the past 12 years[2]. We have, however, considered the possible experimental use of "intelligence" tests in our schools as at present organised and we make the following tentative suggestions in the hope that they may be of some assistance to teachers in schools of various types.

In general we think that the use of tests of "intelligence" as supplementary to the ordinary tests of educational attainments and to standardised scholastic tests is of the greatest value below the stage of the First School Examination. We are of opinion that the use of tests of intelligence by adequately trained teachers and inspectors would chiefiy assist in the following directions:—

(*a*) the proper classification of pupils at entry;

(*b*) the transference of pupils to classes or schools of different types;

(*c*) the detection of the causes retarding or accelerating the development of children; and the adoption, either within the school they attend or at special schools, of the most effective methods of instruction or training in individual cases.

---

[1] See Appendix II.

[2] See Appendix III., pp. 178–183.

By way of preface we should say we are strongly of opinion that it is most important that the "mental ratios" (intelligence quotients) of individual children obtained as a result of the application of "intelligence" tests should as a rule be treated with discretion by the head master or head mistress of the school and should not be shewn except in special circumstances to members of the assistant staff. It is undesirable that the supposed mental ratio of any individual child should be published to the whole staff of the school, and still more that it should be known to the child himself or his fellow pupils.

Further it seems desirable that the "mental ratio" of each individual pupil should be supplemented by notes about the way in which he or she actually attacked the questions put.

We are of opinion that one of the most important uses of the data obtained from such tests is to suggest to the teacher the advisability of reconsidering his judgments about individual pupils which often tend to become stereotyped.

We would also suggest that careful records should be kept both of the results of group and individual tests of "intelligence" and of standardised scholastic tests, on the one hand, and of the after performances of individual pupils, on the other hand, with a view to the comparison of promise of educable capacity as revealed by such tests with the actual achievements of individual pupils in school and after, and also with a view to the elimination of untrustworthy tests of whatever kind and the gradual improvement of testing as a whole.

82. *In Elementary Schools.* (*a*) We are of opinion that in Elementary Schools it might be found useful to apply individual tests[1], and possibly certain group tests, to pupils on entry to the school or senior department at the age of about 7 years, in order to assist the school authorities in the preliminary classification. Individual tests should, however, only be applied when the services of a properly trained person are available. We consider that the data derived from the application of the tests would be useful for pupils of this age in the almost complete absence of other information, e.g., class records of progress in specific subjects and teachers' estimates

[1] Very few group tests of intelligence have as yet been devised suitable for children under 10 years of age. On the other hand, standardised scholastic tests in reading and simple arithmetic would seem to be peculiarly suited for children of about 7 years of age. See § 86.

of intellectual ability, character, powers of memory, attention, and so forth. We think, however, that the data obtained from the use of the tests whether group or individual, should only be regarded as establishing a presumption and should not in the present state of their development be regarded as in any way finally valid.

(*b*) In regard to retarded children or those who are suspected of mental deficiency, we are of opinion that individual tests of intelligence should only be applied by persons who have had adequate training in the technique of applying and marking such tests. Further we entirely approve of the procedure ordinarily adopted at present in cases where a child is to be recommended for transfer to a special school for mentally defective children, whereby the data derived from the application of individual tests of "intelligence" and standardised scholastic tests are always considered in close association with the medical report, the opinion of the child's teachers, the school record, and any available particulars regarding parentage and home conditions. No child should ever be treated as mentally deficient on the ground only of the data afforded by the application of "intelligence" tests, however carefully they may have been administered and marked.

(*c*) We have already expressed our view that it may be desirable that group tests of intelligence should be applied in Elementary Schools to pupils of about 11 years of age who are shortly to be presented for the free place examination, in order that the data thus obtained may be considered in connection with the school records and the teachers' estimate of each individual pupil at the final allocation of free place scholarships after the free place examination[1].

(*d*) We are also of opinion that in doubtful cases where the estimates of the ability and aptitude of individual pupils by different teachers are widely divergent, it would be found helpful to apply individual tests, provided always that they were administered by a person who possessed the necessary training. The "mental ratio" derived from the application of individual tests in such instances might be of use to the head teacher as a check on other information available regarding such pupils.

(*e*) We are disposed to doubt whether it is necessary, or even advisable, to set group tests of intelligence as a part of the

[1] See § 78 (*a*).

ordinary terminal examinations in Elementary Schools, except possibly in those schools where the difficulties of discrimination are exceptionally great. In such cases group tests might be included experimentally in some of the terminal examinations. As a general rule, however, we are disposed to think that it would be sufficient if the data derived from the application of "intelligence" tests to pupils on entrance were checked once or twice before the age of 11 is attained. As we have indicated above[1] group tests might in some instances be applied again before the children are presented for the free place examination.

83. *In Junior Technical Schools.* We are of opinion that experiments might usefully be conducted in the application of group tests to pupils admitted to Junior Technical Schools on the result of the ordinary written, oral and practical examination or of any other tests for entrance as a check on the information thus obtained from the examination and on the opinions formed by teachers of individual pupils. We desire, however, to emphasise our opinion that the data derived from the use of the tests should always be considered in close association with the results obtained from the ordinary written, oral and practical examinations employed in such schools.

84. *In Central Schools.* We understand that the head masters of Central Schools and other types of intermediate schools providing a four years' course for pupils of 11 years and upwards are not infrequently called upon to recommend some of their more gifted pupils for transfer to Secondary Schools at the age of about 14 years. It seems to us that the data afforded by the careful application of group tests of intelligence to such pupils might be of considerable service to the head master as a check on the estimate of the pupil's abilities formed by himself and his staff and the data afforded by the pupil's school record at his Elementary School.

85. *In Secondary Schools.* (*a*) It is well-known that the difficulties of organisation in many Secondary Schools, especially girls' schools, are very great owing to the diversity of attainments in different subjects among the entrants, due to the previous education received by them before entering the Secondary School (whether in Elementary Schools or private schools or through private tuition). We think that the data obtained by

[1] See § 78 (*a*).

the careful application of group tests of intelligence would in some instances be of real value as a partial check on the other data available. The tests might also be used as an aid in classifying entrants of the same attainments.

(*b*) The tests might also be used by head masters and head mistresses of Secondary Schools, as affording additional data to assist them in the task of assigning pupils after one or two years at the school, say at the age of 12 or 13 years, to different divisions of a form.

(*c*) It is well known that reports on pupils in Secondary Schools not infrequently exhibit considerable variation of standard of judgment in regard to the performance and supposed capacity of individual pupils. This probably arises from the fact that particular judgments regarding the aptitude of an individual pupil in the several subjects of the curriculum are sometimes made without any adequate correlation of opinion or standard by different teachers who are specialists in their own subjects. This tendency is, we understand, especially noticeable in the middle forms of some large Secondary Schools, which are often staffed by specialist teachers who give instruction only in their own subject or group of subjects to a considerable number of different classes. The result frequently is that an individual specialist teacher, who is often required to take a number of large classes, must necessarily have a relatively limited knowledge of the idiosyncracies of each individual pupil. We would suggest that in such cases data regarding individual pupils obtained by the application of "intelligence" tests might be of considerable use to the head master or head mistress as a check on the varying judgments of different specialist members of the staff.

PART III.

THE POSSIBLE USE OF STANDARDISED SCHOLASTIC (EDUCATIONAL) TESTS, VOCATIONAL TESTS, PHYSICAL TESTS, TESTS OF MEMORY, PERCEPTION AND ATTENTION, AND TEMPERAMENTAL TESTS, AS AFFORDING DATA BEARING ON EDUCABLE CAPACITY, SUBSIDIARY TO THE INFORMATION YIELDED BY TESTS OF INTELLIGENCE[1]. THE POSSIBLE USE OF STANDARDISED, SCHOLASTIC (EDUCATIONAL) TESTS BASED ON AGE PERFORMANCE AS AN ANCILLARY MEANS OF GAUGING EDUCABLE CAPACITY OF PUPILS IN SCHOOLS OF VARIOUS TYPES.

86. Standardised scholastic (educational) tests, the development of which is described in Sections 33 to 36 of this Report, occupy a position between ordinary examinations and tests of intelligence. Though they may be regarded as a variant of the ordinary oral and written examinations, they resemble "intelligence" tests in that they have been elaborated by actual experiment and statistical evaluation and claim to be based on the average performances of groups of pupils of like age who have enjoyed similar educational opportunities. As has been indicated in Chapter I, a large number of tests of this type in the various school subjects have recently been elaborated in the United States[2]. Most of these American standardised tests of scholastic attainment are, however, unsuited for application to English pupils owing to deep-seated differences in the social and business customs, and in the educational organisation of the two countries[3]. We understand, however, that English psychologists are now devoting considerable attention to the development of satisfactory

---

[1] As we have explained in § 50, we have interpreted the expression "psychological test of educable capacity" in our Reference as including "intelligence" tests, tests of memory, perception and attention, etc., such vocational tests, including tests of manual ability, as are adapted for use in schools and educational institutions, certain physical tests, and tests of such aspects of temperament and character as directly affect educable capacity.

[2] See Appendix VIII for examples of American and English standardised scholastic tests.

[3] See Appendix VI on Grades in American Schools.

tests of this type for English children and we are disposed to think that if adequate tests of this kind were elaborated, especially in the various school subjects, for Elementary Schools and for the lower forms in Secondary Schools up to the age of 14 or 15, they should only be brought gradually and tentatively into use and that the data obtained by their application should in the present state of the development of such tests be regarded as having only a relative value. Further, we think that such tests, when elaborated, would be of use within the school rather than for external examinations. As they would be based on the average abilities, which would have been determined for the various ages or standards in certain fundamental branches of study, they should be useful for discovering the scholastic age of pupils as distinct from their chronological, physiological and mental ages, and for gauging their educational progress. We are of opinion that carefully standardised scholastic tests in simple subjects such as reading and simple arithmetic might prove of real value as an aid to the classification of children in Elementary Schools at about 7 years of age and we suggest that experiments should be made in the use of such tests for this purpose[1].

87. *The possible use of vocational tests, including tests of manual ability, in schools and other educational institutions.* We are disposed to think that certain tests of vocational aptitude, have already reached a promising stage of development and should prove to be of considerable service to teachers, parents, employers, and to public and private bodies engaged in recommending or training children and young persons for certain specific occupations. The data obtained by the use of such vocational tests should always be considered in relation to the data derived from the application of " intelligence " tests and such other information as is available regarding each individual child. Experiments might usefully be conducted in the application of vocational tests in close association with " intelligence " tests to pupils in Elementary Schools and in the lower forms of Secondary Schools, more especially those pupils who appear to have little aptitude for the ordinary school studies. The negative value of such tests might prove to be even greater than their positive value in assisting teachers and parents to dissuade children from entering occupations for which they are shown to be naturally unsuited.

[1] See § 82

On the whole, we are of opinion that the range of these vocational tests in their present state of development is not yet sufficient to warrant us in making any recommendation for their immediate general application. The data afforded by the application of such tests might, however, form an important and useful element in giving vocational guidance to children who are about to leave school, provided always that such data are considered in close relation with school records, the opinion of teachers, any information supplied by the school medical officer, and the performances of the pupils in any external examination, such as the free place examination, for which they may have been presented. Due regard should also, of course, be paid to the personal inclinations of the individual pupil, who should, however, be advised to bear in mind local and general needs and conditions in selecting a career[1].

88. *The possible use of physical tests as an ancillary means of determining educable capacity.* We are of opinion that the data afforded by the application of physical tests in their present state of development are of relatively small value as a subsidiary means of gauging educable capacity. We think that, wherever possible, the data obtained from such tests administered by a trained expert, might form a valuable addendum to the information supplied by the School Medical Officer regarding the general health and physical peculiarities, if any, of each individual pupil. It would seem, however, not improbable that, as a result of further research, these physical tests may be developed and may ultimately yield results which might form a really valuable adjunct to the data derived from the application of " intelligence " tests.

89. *The possible use of tests of memory, perception, attention, etc., as subsidiary to tests of " intelligence."* We are of opinion that, up to the present, tests of memory, perception, attention, imagery and association, have not yet been

---

[1] *Cf.* The following resolution passed by the Second Congress of Psychologists, held at Barcelona in 1921 : " Vocational guidance and selection, besides their scientific, physiological and psychological interest, have an economic and social interest ; they should, therefore, be linked up with agencies for employment or for the relief of unemployment." *International Labour Review*, Vol. V, No. 5 (May, 1922), p. 721.

sufficiently developed to make them of any appreciable value to teachers as supplementing the information afforded by the application of mental tests[1]. It is not improbable, however, that the scope and accuracy of psychological tests of this character may be considerably extended as a result of further research. We are of opinion that if really trustworthy tests of mental activities of a specialised kind, such as memory, perception and attention, could be devised which would at the same time be easy to administer and to mark, such tests might be of considerable use as yielding information regarding individual children that would supplement the data derived from the application of "intelligence" tests.

90. *The possible use of temperamental tests as subsidiary to mental tests.* We are of opinion that no really trustworthy and satisfactory tests of those aspects of temperament which bear directly on educable capacity have yet been elaborated. If satisfactory tests of this type were developed in the future it seems possible that they might be of considerable use as affording important information bearing on educable capacity ancillary to the data derived from the application of "intelligence" tests. For the present, however, we would deprecate the use of any of the existing tests of temperament in schools. In our opinion the only satisfactory method of obtaining any trustworthy data regarding the temperament of individual pupils is the careful collation of reports submitted by competent observers, such as parents and teachers, who have been acquainted with the examinee during a considerable portion of his life, together with such information as can be obtained at a personal interview. In this connection we desire to emphasise the view we have previously expressed that more attention should be devoted to the proper elaboration of the technique of the personal interview. We are disposed to think that something could be done to this end by drawing up questionnaires of facts about temperament to be noted and observed, and possibly also by elaborating rating scales for the registration of such facts in terms of a comparable scheme. On the whole, however, we think that the most important evidence regarding those aspects of character which bear directly on educable capacity is that forthcoming from teachers and others who have had opportunities of observing such traits in an individual pupil over a considerable period of time.

[1] See §§ 16-19.

PART IV.

THE PROVISION OF COURSES FOR TEACHERS IN EDUCATIONAL PSYCHOLOGY, WITH SPECIAL REFERENCE TO THE THEORY AND PRACTICE OF PSYCHOLOGICAL TESTS, AND THE QUESTION OF TAKING STEPS TO CO-ORDINATE AND ORGANISE RESEARCH AND TO DISTRIBUTE INFORMATION BEARING ON THE APPLICATION OF PSYCHOLOGICAL AND STATISTICAL METHODS TO EDUCATION.

91. *The question of the training which should be given to teachers, school doctors, and other persons interested in education to render them competent to administer group and individual tests of intelligence respectively.* We are of opinion that the devising, standardising, and final interpretation both of individual tests and group tests, and of their methods of application should be entrusted only to recognised psychological experts. But we think that the satisfactory application of group tests which are set in the form of a written examination paper, demands relatively little special training, though it seems most important that the instructions given to the supervisor who distributes the papers to the examinees should be exactly carried out. The technique is relatively simple, and a teacher should be able to acquire the necessary knowledge by attending a short course on educational psychology, with special reference to the significance of group tests of " intelligence " and standardised scholastic tests to the technique of administering and marking them. We accordingly suggest that local education authorities should take steps and to organise short courses in educational psychology carefully restrained in scope and with special reference to the technique of applying and marking group tests of " intelligence " and standardised scholastic tests for teachers already employed and for those in training in order to give them an appreciation of the significance of tests and to impress on them the need for exact method in their application. We think that teachers should, after a short course of this kind, be able to mark group tests and to interpret the results in reference to the pupils in their own schools, though it would be hazardous if they were to attempt to interpret the results generally in reference to children of the same age.

On the other hand, we are of opinion that for the satisfactory application by teachers, school medical officers and others of individual "intelligence" tests, and for the accurate interpretation of the data obtained by their use, a training in experimental psychology, in the technique of applying the tests, and in the use of statistical methods, is indispensable. The results obtained by the application of individual tests of "intelligence" by untrained persons should be received with the utmost caution, and are, in our judgment, almost devoid of value.

We accordingly suggest that longer courses in experimental psychology, with particular reference to the theory of "intelligence" tests and the technique of applying, marking and interpreting both individual tests and group tests, should be organised by Universities and University Colleges acting in concert with local education authorities, for experienced teachers and for those who are taking a deferred third-year course, or a one-year professional course. Such courses should always include some instruction in the elements of statistical methods.

We are of opinion that when psychological tests are used to ascertain whether children are mentally deficient, they should be applied only by persons possessing an adequate knowledge of the technique of applying, marking, and interpreting individual tests of intelligence and standardised scholastic tests. We accordingly recommend that the Board should take appropriate steps to ensure this, either by supplementing the conditions under which they at present approve of school medical officers whose duties involve the examination of children suspected of mental deficiency, or by otherwise ensuring that an experienced expert is available for this purpose.

92 *The question of provision for the co-ordination and correlation of the work of psychologists, teachers and statisticians in regard to the application of psychological and statistical methods to education.* On a review of the available evidence we are convinced that it is important that existing psychological tests should be kept constantly under review, in order to increase the accuracy of the methods employed, to enlarge their scope and to render the data afforded by them more valid. In particular, it seems most desirable that psychologists, in collaboration with teachers who have had some training in psychology, should elaborate new and alternative tests and revise the methods of

application. Further, it seems to us highly desirable that all available information regarding the latest developments of the various types of psychological tests of educable capacity should be collected by some central organisation and placed at the disposal of psychologists, teachers, administrators and other persons interested in education. For these purposes further scientific research is required. It seems, therefore, desirable that the Board of Education should take appropriate steps, possibly in association with other Government Departments interested directly or indirectly in the development of psychological tests, to appoint a permanent advisory committee to work in concert with the Psychological Departments of the English and Welsh Universities and other organisations engaged in the work of research.

Such a permanent advisory committee might well include administrative officers and inspectors of the Board and other Government Departments interested in tests, and of Local Education Authorities, teachers, medical officers, psychologists and trained statisticians. We think, further, that Local Education Authorities should be encouraged in cases where financial conditions permit, to appoint whole or part-time psychological advisers to collaborate with the school medical officers and the teachers in schools of various types, and generally to supervise the application of tests and other psychological methods bearing on education in the area. In this connection we desire to reiterate our opinion that psychological tests should be regarded as being only the most prominent of several ways in which a knowledge of psychology and statistical methods is being brought to bear on education.

PART V.

CONCLUSIONS AND RECOMMENDATIONS.

93. Our conclusions and recommendations are as follows :—

The Various Types of Psychological Tests of Educable Capacity.

(1) That, while the terminology employed in any discussion of psychological tests of educable capacity in their present state of development is necessarily provisional, the expression

"psychological test of educable capacity" may for purposes of convenience be interpreted as including:—

(i) Tests of "intelligence," *i.e.* tests designed to measure that general ability which is held by many psychologists to underlie the various special activities of the mind.

(ii) Standardised scholastic tests, *i.e.* tests of attainments in particular school subjects, such as reading and arithmetic, elaborated by actual experiment and statistical evaluation*.

(iii) Such vocational tests, including tests of manual ability[1], as are suitable for application in schools and educational institutions.

(iv) Tests of mental activities of a specialised kind, *e.g.*, tests of memory, perception, attention, imagery, association.

(v) Certain physical tests which have been suggested as a means of assessing educable capacity.

(vi) Tests of such aspects of temperament and character as bear directly on educable capacity. (§ 50).

(2) That the distinctions involved in the above classification are themselves founded on hypotheses, and, however convenient for purposes of analysis, should not be interpreted as if they were finally valid; in other words, these distinctions are probably best regarded as first approximations to the truth, and as such are of considerable value for working purposes, but possess only provisional validity. (§ 50).

(3) That the devising, standardising and final interpretation of all types of psychological tests thould be entrusted only to recognised experts, who should take counsel with those who are in close touch with school life. (§ 91).

(4) That up to the present the only kinds of psychological tests of educable capacity that have been sufficiently developed to be of much service in schools for the purpose of diagnosing and assessing such capacity are tests of "intelligence," standardised scholastic tests, and to a less extent vocational tests. (§§ 81–87).

Tests of "Intelligence."

(5) That from one point of view tests of "intelligence" represent a considered attempt to apply methods of quantitative and qualitative evaluation on psychological lines to groups of

* See footnote on p. 145.

[1] See footnote [1] on p. 62.

children and adults who are assumed to have been subjected to a similar general environment and like conditions of life. The tests attempt to gauge "intelligence" in relation to elementary forms of acquired knowledge. They are to be regarded as an attempt to apply a knowledge of psychology and statistical methods to examinations intended primarily to discover ability rather than attainments. (§§ 55–56, 61, 62, 75–78, 82–85).

(6) That though in the present state of development of "intelligence" tests it is hardly possible to explain simply and precisely what is measured or tested by them, it is tolerably well established that these tests, when properly constructed, applied and interpreted, have shown themselves capable of giving a useful common measurement of what teachers generally call capacity or intelligence. They indicate how far a child is capable of learning provided he is not prevented by extraneous circumstances or faults of temperament from making proper use of the opportunities offered him. (§ 55).

(7) That, though there is considerable difference of opinion among psychologists regarding hypotheses about the nature of general "intelligence," there is nevertheless a large measure of agreement in regard to the connotation of the phrase for practical purposes. It appears to be a general mental ability operating in many different ways, given as part of the child's natural endowment, as distinct from knowledge or skill acquired through teaching or experience, and more concerned with analysing and co-ordinating the data of experience than with mere passive reception of them. As, however, this term "intelligence" is employed by psychologists in a technical sense, it seems desirable to place it in inverted commas, in order to indicate that it is used with a special meaning. (§§ 53, 54).

(8) That the main presuppositions underlying the use of "intelligence" tests appear to be as follows:—

(i) That there are certain mental factors which remain more or less constant during the lifetime of individual human beings:

(ii) That methods of examination have been discovered or can be discovered by which these factors in any individual can to a great extent be ascertained and differentiated from the results of training and ducation.

These are, however, subject to the important condition that, in order that the tests may yield valid results, the persons tested must not be drawn from environments that are widely dissimilar nor have been subjected to widely dissimilar conditions of life. (See No. 5). (§ 56).

(9) That there is general agreement that tests of "intelligence" are of value as supplements to, but not as substitutes for, the present methods of estimating individual capacity: that this supplementary testing is of greatest value below the stage of the First School Examination, but that its value must inevitably be reduced so long as it is restricted by the limitations attaching to any system of examination which finally determines so early as the age of 11 the opportunities of a child to proceed to places of higher education. (§§ 72–78, 81–85).

Group and Individual "Intelligence" Tests.

(10) That group tests are more suitable for use in schools and in external examinations than individual tests since the latter take much longer to apply and demand a special training for their proper application. (§§ 61, 62, 78, 82–85).

(11) That up to the present very few group tests have been constructed which are suitable for application to children under 10 years of age.

That the individual tests devised up to the present are especially useful for young children under about 10 years of age, but are of comparatively little use for older children, particularly those over the age of 15 or 16 On the other hand, group tests have been devised which are suitable both for children from 10 to 16 and upwards and for adults. (§§ 62, 82).

(12) That the satisfactory application of group tests, which have been properly elaborated and standardised and are set in the form of written papers, demands relatively little special training, though it is most important that the instructions given to the supervisor who actually conducts the test should be carried out with precision. The technique is relatively simple, and a teacher should be able to acquire the necessary knowledge by attending a short course on educational psychology with special reference to the significance of group tests and to the technique of administering, marking and interpreting them. (*Cf.* No. 15). (§§ 70, 91).

(13) That for the satisfactory application by teachers, School Medical Officers and others of individual tests of intelligence, and for the accurate interpretation of the data obtained by their use, a careful training in psychology, in the technique of applying the tests, and in the use of statistical methods is indispensable. Results obtained from the application of individual tests by untrained persons should be received with the utmost caution and are, in our opinion, almost devoid of value. (§§ 57, 91).

(14) That, when individual tests of " intelligence " are applied to subnormal children with a view to ascertaining whether they are mentally deficient, the data derived from the application of such tests should never be regarded as finally valid in themselves, and should always be considered in close association with the medical report, the opinion of the child's teachers, the school record and any available particulars regarding parentage, home conditions and general environment. In particular, no child should ever be treated as mentally deficient solely on the evidence afforded by the application of " intelligence " tests, or of standardised scholastic tests, however carefully they may have been administered and marked. (§§ 58, 82(b)).

The Provision of Courses in Experimental Psychology and in the Elements of Modern Statistical Methods designed to qualify Teachers and Medical Officers to use Group and Individual " Intelligence " Tests respectively.

(15) That it is desirable that steps should be taken by local education authorities to organise short courses in educational psychology with special reference to the theory of group intelligence tests and the technique of applying and marking them for teachers and for persons engaged in the work of education, in order to give them an appreciation of the significance of such tests and to impress upon them the need for method in their application. (*Cf.* No. 12). (§§ 70, 91).

(16) That it is desirable that longer courses in experimental psychology, with particular reference to the theory of " intelligence " tests and the technique of applying, marking and interpreting both individual and group " intelligence " tests, should be organised by Universities and University Colleges (acting in concert with local education authorities) for experienced teachers and for those who are taking a deferred third-year course or a one-year professional course.

Such courses should include some instruction in the elements of statistical methods. (*Cf.* No. 13). (§§ 70, 91).

(17) That it is important that when psychological tests are used to ascertain whether children are mentally deficient they should be applied only by persons (whether School Medical Officers or others) possessing an adequate knowledge of the technique of applying, marking and interpreting individual "intelligence" tests and standardised scholastic tests.

We accordingly recommend that the Board should take appropriate steps to ensure this, either by supplementing the conditions under which they at present approve of School Medical Officers whose duties involve the examination of children suspected of mental deficiency or by otherwise ensuring that a recognised expert is available for this purpose. (§§ 82(b), 91).

Standardised Scholastic Tests.

(18) That it is desirable that carefully devised experiments should be made in the application of standardised scholastic (educational) tests suitable for use in English schools with a view to determining how far they afford trustworthy subsidiary data bearing on educable capacity. (§ 86).

(19) That standardised scholastic tests in simple subjects, such as reading and the easier processes of arithmetic, might probably be used with good results by head teachers of Elementary Schools to assist them in classifying pupils entering at about the age of 7. It is advisable that experiments should be conducted in the use of standardised scholastic tests for this purpose; up to the present very few group "intelligence" tests have been devised suitable for children under 10, and the satisfactory application of individual "intelligence" tests, especially to young children, demands special training on the part of the teacher. (*Cf.* No. 13). (§§ 62, 66, 86, 92).

The Application of Psychological Knowledge to Ordinary Examinations and the Possible Use of "Intelligence" Tests as Adjuncts to Certain Public Examinations for Young Children.

(20) That the new "psychological" tests may probably exercise an important influence on public examinations of the ordinary type. It seems probable that the whole theory and

technique of ordinary examinations will gradually be considerably modified by applying to them, so far as possible, the principles adopted in constructing and standardising tests of " intelligence " and standardised scholastic tests. (§§ 49, 51, 52, 66, 73, 77, 86).

(21) That, if effective steps were taken to apply a knowledge of psychology and of the new technique of testing to written examinations of the ordinary type (more especially those intended for younger pupils, such as the Free Place Examination), the case for using group tests of " intelligence " as adjuncts to such examinations would be *pro tanto* weakened. In other words, that if questions in written examinations for younger children were always set with due regard to the peculiarities of the child mind, both in the form and matter of the questions and in their arrangement in the written papers, and if the same scientific methods were employed as in " intelligence " tests and standardised scholastic tests, such examinations would prove a more effective means of discovering ability in young children than those now in use, even apart from the application of group tests of " intelligence." (§§ 77, 78.)

(22) That the considerations stated above apply with even greater force to oral examinations, especially those for younger children. In our opinion, one of the most promising lines of advance in the whole range of examinations designed primarily to discover ability is to be found in a careful and detailed study of the legitimate aims and inevitable limitations of *vivâ voce* examinations and their technique. We accordingly suggest that psychologists should be asked to make a special study of the possibilities of oral interviews and oral examinations. (§§ 63, 79, 80.)

(23) That the value of " intelligence " tests in selecting younger children for free places, for entrance to Secondary Schools, and for admission to Central Schools should be investigated by tentatively adding group tests to the customary written examinations ; and that the relative merits of the two sections of the examinations should be estimated by calculating the correlation between the separate results and the subsequent development of the pupils. (§§ 78, 82.)

(24) That further research should be instituted into the use of individual " intelligence " tests in connection with oral examinations for all ages or types. (§§ 79, 80.)

(25) That the current individual " intelligence " tests, whether they be the Binet-Simon Tests or variants of the Binet Scale, should be somewhat modified when applied to rural children or to children living in districts where there is a strong local dialect. We are of opinion that these tests, having been designed primarily for urban children, cannot properly be applied in all instances to rural children without such variations as may be needed to avoid possible misunderstanding. (§ 57).

The Possible Uses of Tests of " Intelligence " within Schools of Different Types.

(26) That experiments might usefully be made in the application of group " intelligence " tests to pupils in Elementary, Central, Junior Technical and Secondary Schools with a view to obtaining data which would assist the school authorities in the classification and promotion of pupils, and in making recommendations for the transfer of individual children to other types of school.

It is desirable that careful records should be kept both of the results of group and individual tests of " intelligence," on the one hand, and of the after performances of individual pupils, on the other hand, with a view to the comparison of promise of " intelligence " as revealed by the tests with the actual achievements of individual pupils in school and after, and also with a view to the elimination of untrustworthy tests of whatever kind and the gradual improvement of testing as a whole. (§§ 81–85).

(27) That the data afforded by the use of " intelligence " tests should not be regarded as possessing any final validity, and should always be considered in close association with the whole body of information regarding individual pupils available from other sources. (§ 81).

(28) That teachers who use intelligence tests should realise that the so-called " mental ratios " (intelligence quotients) of individual children obtained by the application of such tests represent a succinct and highly abstract method of presenting the results, and that the mental ratio of any individual child should always be used with discretion and in association with the information available from other sources. It seems undesirable that the mental defects of individual children as revealed

by low mental ratios should be made known to the whole staff of the school, still less to such children themselves and their fellow pupils. Further, it is advisable that the mental ratio of each individual pupil should always be supplemented by notes about the way in which he or she actually attacked the questions set. (§§ 59, 81).

(29) That one of the most important uses of the data obtained from such tests is to make the teacher reconsider his estimates of individual pupils, which often tend to become stereotyped. A pupil's success or failure in the tests would probably suggest to the teacher the desirability of modifying his judgments of the mental powers of individual pupils and possibly of changing his methods. (§ 81).

Vocational Tests.

(30) That certain tests of vocational aptitude have already reached a promising stage of development, and should prove to be of considerable service to parents and employers and to public and private bodies engaged in recommending or training children and young persons for certain occupations, and still more for discouraging children from entering occupations for which they are unsuited ; but that the range of such tests is not yet sufficient for any recommendation to be made for their immediate general application. (§§ 67, 87).

(31) That in cases where vocational tests and tests of manual ability are experimentally applied to children still at school, the data obtained thereby should, as in the case of " intelligence " tests, always be considered in association with the information obtainable from other sources, *e.g.*, the data afforded by the use of " intelligence " tests, school records, the opinion of teachers, medical data, and any particulars regarding parentage and environment. (§ 87).

(32) That in order to facilitate the advising of boys and girls as to the type of occupations for which they are fitted, more information than that afforded by examination results and tests of " intelligence " and of vocational aptitude is needed ; and that for the purpose of securing more complete data on which advice may be based, all Committees and bodies dealing with

the placing of children in employment should be encouraged to consult teachers and members of the School Medical Service as well as psychologists. (§ 87).

Psychological Tests of Educable Capacity other than "Intelligence" and Vocational Tests and Standardised Scholastic Tests.

(33) That tests of memory, perception, attention, etc., in their present state of development, are of little use to teachers as a means of acquiring information subsidiary to that derived from the application of "intelligence" tests (§§ 18, 19, 49, 89).

(34) That physical tests, in their present state of development, are of very little use to teachers for the purpose of affording data ancillary to the information derived from the application of "intelligence" tests (§§ 6–8, 68, 88).

(35) That tests of temperament and character, in their present state of development, are practically useless to teachers for the purpose of affording trustworthy information on such aspects of temperament and character as bear directly on educable capacity. (§§ 43, 49, 64, 90).

Provision for Co-ordinating and Rendering Available for General Use the Results of Scientific Research in Regard to Psychological Tests of Educable Capacity.

(36) That it is desirable that existing psychological tests, and more especially intelligence and vocational tests and standardised scholastic tests, should be kept constantly under review, in order to increase the accuracy of the methods employed, to widen their scope, to render their conclusions more valid, and to reduce any danger of special preparation beforehand, and that in particular new and alternative tests should constantly be instituted and methods of use revised. (§§ 71, 92).

(37) That for these purposes it is advisable to encourage scientific research on the subject, and to render available for use in schools and elsewhere the results of such research, and we accordingly suggest that the Board of Education, either acting independently or in association with other Government Departments interested in psychological tests, should set up an Advisory

Committee to work in concert with University Departments of Psychology and other organisations engaged in the work of research. (§§ 71, 92).

(Signed) W. H. HADOW (*Chairman*).
P. ABBOTT.
*GEORGE ADAMI.
S. O. ANDREW.
ERNEST BARKER.
E. R. CONWAY.
D. H. S. CRANAGE.
GORELL.
IVOR H. GWYNNE.
F. HAWTREY.
P. R. JACKSON.
STANLEY LEATHES.
A. J. MUNDELLA.
BERTHA S. PHILLPOTTS.
ROBERT H. PICKARD.
FRANK ROSCOE.
R. P. SCOTT.
E. M. TANNER.
R. H. TAWNEY.
W. W. VAUGHAN.
J. A. WHITE.

ROBERT F. YOUNG (*Secretary*).
27th March, 1924.

* Dr. Adami objects to the inclusion of standardised scholastic tests under the head of "psychological tests of educable capacity."

APPENDIX I.

(A) *List of Witnesses.*

Miss G. M. ANDERSON, Head Mistress of the Girls' Department of the Ben Jonson Council School, Harford Street, Stepney, London.

Mr. P. B. BALLARD, D.Litt., District Inspector under the Education Department of the London County Council.

Mr. ANDREW BELL, Head Master of the Erith County Secondary School.

Mr. WILLIAM BROWN, M.D., D.Sc., M.R.C.P., late Reader in Psychology in the University of London, Reader in Mental Philosophy in the University of Oxford.

Mr. CYRIL BURT, D.Sc., Psychologist to the London County Council.

Mr. E. G. CLARKE, Assistant Master at the L.C.C. School, Star Road, West Kensington, London.

Mr. CHRISTOPHER COOKSON, Secretary to the Secondary School Examinations Council.

Miss W. M. CROSTHWAITE, Head M stress of the Colchester County School for Girls.

Mr. R. R. DOBSON, Head Master of Pate's Grammar School for Boys, Cheltenham.

Mr. W. H. A. DOCKERILL, Head Master of the Kenmont Gardens L.C.C. School, College Park, London.

Mr. G. B. DODDS, Head Master of the Harmood Street L.C.C. School for the Mentally Defective, Chalk Farm, London.

Miss BEATRICE EDGELL, Ph.D., University Reader in Psychology in the University of London.

Mr. A. EICHHOLZ, C.B.E., M.D., Chief Medical Inspector under the Board of Education.

Mr. M. GOMPERTZ, Head Master of the Leyton County High School for Boys.

Mr. HUGH GORDON, formerly H.M. Inspector of Schools.

Miss F. R. GRAY, High Mistress of St. Paul's Girls' School, Hammersmith, London.

The late Mr. J. A. GREEN, Professor of Education at the University of Sheffield.

Mr. C. H. GREENE, Head Master of Berkhamsted Grammar School.

Mr. F. T. HOWARD, H.M. Staff Inspector of Elementary Schools.

Mr. G. JONES, Head Master of the Tennyson Street L.C.C. School, Battersea, London.

Mr. M. W. KEATINGE, D.Sc., Reader in Education, University of Oxford.

Mr. C. W. KIMMINS, D.Sc., late Chief Inspector in the Education Department of the London County Council.

Mr. E. O. LEWIS, D.Sc., M.R.C.S.

Mr. B. MUSCIO, D.Sc., Professor of Philosophy at the University of Sydney, N.S.W., formerly Investigator to the Industrial Fatigue Research Board.

Mr. C. S. Myers, C.B.E., M.D., Sc.D., F.R.S., Director of the National Institute of Industrial Psychology.
Mr. T. Percy Nunn, D.Sc., Principal of the London Day Training College, and Professor of Education in the University of London.
Mr. T. H. Pear, Professor of Psychology in the Victoria University of Manchester.
Mr. C. A. Richardson, H.M. Inspector of Schools.
Mr. F. C. Shrubsall, M.D., Principal Assistant Medical Officer under the London County Council.
Mr. Charles E. Spearman, Ph.D., F.R.S., Grote Professor of Philosophy of Mind and Logic, in the University of London.
Mr. Max Tagg, Head of the Engineering Department at the Acton and Chiswick Polytechnic and Head Master of the Junior Technical School.
Mr. Godfrey H. Thomson, D.Sc., Ph.D., Professor of Education at Armstrong College, University of Durham.
Mr. J. G. Tibbey, Head Master of the Mile End Central School, Myrdle Street, Stepney, London, and Editor of "Educational Research."
Mr. Graham Wallas, late Professor of Political Science in the University of London.
Mr. Frank Watts, Assistant Inspector of Schools, Board of Education, late Lecturer in Psychology in the University of Manchester and in the Department of Industrial Administration of the College of Technology, Manchester.
Miss Mary T. Whitley, Ph.D., Professor at Teachers College, Columbia University, New York.
Mr. W. H. Winch, District Inspector of Schools under the London County Council.

(B) *List of Persons and Organisations who sent Memoranda to the Committee.*

Mr. H. G. Abel, Head Master of St. Olave's and St. Saviour's Grammar School for Boys, Bermondsey, London.
Blackpool Education Committee.
Mnr. S. C. Bokhorst, Ph.D., Amsterdam.
Bradford Education Committee.
Lt.-Col. Sylvester Bradley, Medical Inspector of Recruits, Eastern Command.
Miss M. D. Brock, Litt.D., Head Mistress of the Mary Datchelor Girls' School, Camberwell, London.
Mr. W. A. Brockington, O.B.E., Director of Education for Leicestershire.
Mnr. N. Brugmans, Ph.D., Director of the Dr. D. Bos Foundation, Groningen.
Mr. Arthur Burrell, late Principal of the Borough Road Training College, Isleworth.

Señor José Castillejo, Director of the Instituto Escuela de Segunda Eseñanza, Madrid.
The Civil Service Commissioners.
Mr. R. F. Cholmeley, Head Master of Dame Alice Owen's Boys' School, Islington, London.
Monsieur A. G. Christiaens, Director of L'office Intercommunal pour l' Orientation professionelle, Brussels.
Monsieur E. Claparède, President du Conseil de l'École des Sciences de l'Education (Institut J. J. Rousseau), Geneva.
Mr. A. C. Coffin, late Director of Education for Bradford.
Rev. E. Dale, D.Litt., Head Master of the Latymer Upper School, Hammersmith, London.
Mr. James Drever, D.Phil., Combe Lecturer in Psychology at the University of Edinburgh.
Wing Commander Flack, C.B.E., of the Medical Department of the Royal Air Force.
Pan V. Forster, Ph.D., Director of the Institute of Applied Psychology at the Masaryk Academy of Labour, Prague.
Mr. D. Kennedy Fraser, Psychological Adviser to the Glasgow Education Authority.
Rev. F. H. George, Head Master of Bloxham School, Banbury.
Mr. James Graham, Director of Education for Leeds.
Mr. J. H. Hallam, Chief Officer for Higher Education (Secondary) for the West Riding of Yorkshire.
Miss Victoria Hazlitt, Lecturer in Psychology at Bedford College for Women, University of London.
Mr. William A. F. Hepburn, Head Master of Newburn Hall Council School, Lemington-on-Tyne (late Senior Assistant to Dr. Drever in the Psychological Laboratory at Edinburgh University).
Mr. A. E. Ikin, LL.D., Director of Education for Blackpool.
Mrs. S. S. Isaacs.
Mr. G. A. Jaederholm, Professor of Philosophy at Gothenburg College (Högskola).
Miss A. Kalshoven, Ph.D., Amsterdam.
Mr. F. J. Kemp, Head Master of the Haberdashers' Aske's Hampstead Boys' School, London.
Herr Otto Lipmann, Ph.D., Director of the Institute of Applied Psychology at Berlin.
Leeds Education Committee.
London Education Committee.
Miss Margaret McFarlane, Bedford College for Women, London.
Miss Helen M. Madeley, Assistant Director of Higher Education for Warwickshire.
Pan J. Mauer, Ph.D., Librarian of the Comenius Paedagogical Institute at Prague.
Mr. George H. Miles, D.Sc., Assistant Director and Secretary of the National Institute of Industrial Psychology.
Mr. Edward J. W. Mountford.

Mr. A. A. Mumford, M.D., Medical Officer of Manchester Grammar School.
Mr. Kai Mygind, Danish Chamber of Manufacturers, Copenhagen.
The National Institute of Industrial Psychology.
Northumberland Education Committee.
Mr. Hyde Pickford, Acting Director of Education for Bradford.
Monsieur Henri Pieron, Director of the Laboratoire de Psychologie Physiologique de la Sorbonne, Université de Paris.
Mr. Rudolf Pintner, Ph.D., Professor at Teachers College, Columbia University, New York.
Mr. Martin L. Reymont, Ph.D., Christiania.
Mr. Robert R. Rusk, Ph.D., Principal Lecturer in Education to the St. Andrew's Provincial Committee for the Training of Teachers.
Signor Sante de Sanctis, D.Phil., Professor of Psychology at the Royal University of Rome.
Mr. Frank Sherwood, Leeds.
Mr. George Smith, Master of Dulwich College.
Herr W. Stern, D.Phil., Professor of Philosophy at the University of Hamburg.
The Professors of the Psychological Faculty of Teachers College, Columbia University, New York.
Mr. Edward L. Thorndike, Ph.D., Professor at Teachers College, Columbia University, New York.
Mr. T. P. Tomlinson, Head Master of Love Lane Council School, Pontefract.
Warwickshire Education Committee.
Mr. Henry J. Watt, D.Phil., Lecturer on Psychology in the University of Glasgow.
Mnr. van Wayenburg, Ph.D., Director of the Psycho-Technical Laboratory, Amsterdam.
Mr. C. Williams, O.B.E., Director of Education for Northumberland.
West Riding of Yorkshire Education Committee.

APPENDIX II.

Short Accounts of some Experiments recently conducted in England in the use of Group Tests and Individual Tests in Free Place Examinations and in Schools of Different Types.

(1) *Experiments in the Use of* (*a*) *Group Tests of Intelligence, and* (*b*) *Standardised Scholastic Tests at Blackpool* (*summarised from data supplied by the Director of Education*).

(*a*) The selection of candidates for scholarships to Central and Secondary Schools in the County Borough of Blackpool an 1922 was made by the following stages : in the first instance, an examination or review of the work of all pupils in Elementary Schools

between the ages of 11 and 12 was made. About 900 children sat for the preliminary examination, and about half this number were selected to sit for a further competitive examination in English and arithmetic. The marks gained at the written examination and at the oral examination, together with a report from the head teacher of the Elementary School which the candidate had attended, were taken into consideration in the award of scholarships. As a result 30 candidates were sent to Secondary Schools and 240 to Central Schools. As an experiment one of the published series of group intelligence tests was given to all the candidates selected for the second examination. The papers containing these tests were carefully marked, and the intelligence coefficient of each child was calculated from them, and a list made in order of merit. It was found that the list was not identical with the scholarship list, and it was noted that 16 of the first 30 children on the intelligence tests list were included in the first 30 on the scholarship list. This result, however, did not afford material for deciding whether the intelligence test or the ordinary examination afforded the more trustworthy result. Accordingly, as a further experiment, 92 of the first 100 pupils on the scholarship list were set four other group intelligence tests. The papers thus compared were three American tests and two English tests. The following table gives the position of each list for certain of the pupils :—

Junior Scholarship List.	*Test No.1.*	*Test No.2.*	*Test No.3.*	*Test No.4.*	*Test No.5.*
1	21	1	3	13	3
10	7	4	4	3	2
20	36	13	71	31	43
30	2	7	67	80	63
40	63	55	57	37	66
50	25	41	41	4	48
60	56	83	85	80	85
70	80	52	29	43	71
80	8	27	35	52	33
90	77	81	60	88	41
92	5	87	69	84	89

Dr. Cyril Burt worked out the correlation of these results with the scholarship results, arriving at the following coefficients respectively : ·35, ·55, ·51, ·46, ·50, with a probable error from ·05 to ·06. Dr. Ikin believed that some guidance could be obtained from the use of group tests which might be of great value in deciding borderline cases, but the chief result of the experiment in his opinion, was that though intelligence tests of this kind were being

developed on right lines, they had not yet achieved such a state of perfection as would justify the award of scholarships on group intelligence tests alone. Dr. Ikin had found that a comparison of the work of the pupils subsequent to the application of the tests was showing that these tests did assist in the discovery of educable capacity, and he believed that if a group test were given in addition to the ordinary attainment test, there would be a greater probability that the more suitable candidates would be selected for higher education, than if either type of test were employed alone.

Dr. Ikin states that he is continuing his experiments with group tests of intelligence ; and that it is probable that this year he will be able to combine a group intelligence test with the ordinary scholastic examination for awarding scholarships.

(*b*) Dr. Ikin has recently applied standardised tests of arithmetic to all the children aged 7 to 14 in the ordinary Elementary Schools of Blackpool. The tests were applied after the scholarship examination had been set, and immediately after the children so selected had been transferred to Secondary or Central Schools. The object was in part to discover whether any children of super-normal ability in arithmetic had been missed by the scholarship examination.

Those used were the tests of mechanical and problem arithmetic taken from Dr. Cyril Burt's *Mental and Scholastic Tests.* They were set and marked by the teachers ; the tabulated results were submitted to Dr. Burt, with a request for comments ; and his reply was then circulated to all the teachers concerned.

The following are some of the conclusions drawn from the experiment :—

(i) The use of such standardised scholastic tests in the hands of the ordinary teacher presents no difficulties, and yields suggestive information in regard to the level of attainments of the children in the schools within a district.

(ii) Tests standardised in one urban area, such as London, seem directly applicable, with little or no change, to another urban area, such as Blackpool.

(iii) Dr. Burt states that, as compared with London children, the younger children at Blackpool, particularly those aged 7, did unusually well in the tests ; and that, in general, the Blackpool children were better at problem work than they were at mechanical work. It is interesting to note that the level of attainments in two such widely separated districts is all but identical.

(iv) The relative inferiority of the pupils over the age of 11 appears, in Dr. Ikin's view, to confirm the anticipation that the best pupils had already been successfully selected by the annual scholarship examination.

(2) *Results Obtained by the Use of Group Tests of Intelligence in the Examination for Free Places and for Entrance to Secondary and Central Schools at Bradford in* 1922 (*summarised from data supplied by the Director of Education*).

In the examination for free places and for admission to Secondary Schools and Central Schools, held at Bradford in March, 1922, the maximum mark for group intelligence tests was 100, while that for arithmetic and English combined was 200. The results obtained by the application of the intelligence tests were not to be taken into account, except in border line cases. As an instance of the effect of the inclusion of the marks for intelligence tests, it was pointed out that 448 pupils would thereby be raised from between 33 per cent. and 50 per cent. to over 50 per cent. It was found that pupils admitted on these tests justified their choice by profiting from a secondary education. It was stated that a number of cases had already been investigated, and, although the total number of candidates so examined had been small, nevertheless the enquiry, so far as it had gone, had confirmed the value of group intelligence tests as an indication of ability to profit by a Secondary School education. The number of cases investigated in the Secondary Schools was 48 boys and 44 girls, and it was found that in general the results obtained from the intelligence tests corresponded with the progress made in the Secondary School, as indicated by the terminal examinations, to a rather greater extent than obtained when a comparison was instituted between the marks awarded in the ordinary scholarship examination and the marks given in the Secondary Schools.

(3) *The Use of Intelligence Tests in Examinations for Free Places offered by the Leicestershire Education Committee* (*summarised from a memorandum sent by the Director of Education*).

All pupils in the Elementary Schools in Leicestershire (with the exception of those not recommended by their head teachers) between the ages of 11 and 12 on the 1st June in any year are examined in English and arithmetic, and also undergo an oral examination. For the marking of papers in the written examination, and for the conduct of the oral examination, district boards consisting entirely of Secondary and Elementary School teachers are appointed.

The term "oral examination" is used to cover a consultation between the two visiting teachers (one secondary and one elementary) and the head teacher of the Elementary School as to the capacity and promise shown by the pupils. For purposes of standardisation the oral examiners are supplied privately with a "provisional qualifying mark" based on the results of the written examination in English and arithmetic. A preliminary order of qualification, drawn up by the head teacher, enables the examiners to make ready reference to any pupils who may have failed to do themselves justice in the written papers.

The visiting examiners may question the pupils on errors occurring in their written papers and may test their knowledge in other subjects, especially history and geography. In addition, the visiting teachers may impose any reasonable mental tests of general capacity.

For the purpose of estimating the actual intelligence, as distinct from the acquired knowledge of the children, the examiners, generally speaking, had not relied solely on mental tests according to the text-book models, but they had taken into consideration the results of the written papers, and especially the opinion of the head teacher and the record of school work.

The total impression left upon the minds of the examiners was the correctness of the head teachers' estimates, subject to the invaluable check (which a general examination over a large area provided) upon possible faults of judgment and false standards of knowledge.

Experience did not show whether mental tests would have resulted in the same order of placing the pupils, had the examiners approached their task with an open mind and with no previous knowledge of the capacity and attainment of the pupil. Some of the most experienced teachers who had conducted oral examinations in Leicestershire were of opinion that "intelligence" tests, whether oral or written, should not be the sole means of estimating the quality of candidates for admission to Secondary Schools. Such tests indeed revealed presence of mind and quickness in "seeing the point," but ignored application and staying power. Nevertheless, they were invaluable in the oral examination.

(4) *The Use of Group Tests of Intelligence in Examinations for Free Places offered by the Northumberland Education Committee (summarised from memoranda sent by (a) Professor Godfrey H. Thomson, and (b) the Director of Education.*

(*a*) In a memorandum sent to the Committee by Professor Godfrey Thomson, he stated he he had tested about 3,000 children in Northumberland in 1921 and about 14,000 in 1922 by means of group tests devised by him for the purpose, as well as some 2,000 children in other parts of Great Britain during the preliminary process of standardising the tests. He was convinced of their value in giving due weight to inborn ability in the exceedingly important and difficult task of selecting children for different types of schools at the age of 11. He was well aware, from his own work as well as from his study of the publications of Burt, Gordon and others, that intelligence tests, especially group tests, were by no means uninfluenced by schooling, so that a child from a poor school or from an uncultured home was still heavily handicapped. But he believed that this handicap was far less than the handicap felt by such a child in attempting, say, a County Minor Scholarship examination. He had re-tested a small number of children for several

years by individual tests, and had found constant intelligence quotients. Further, he had re-tested 11,000 children at a year's interval by his own group tests, finding practically the same intelligence quotient in 85 per cent. of the cases, but changes ranging up to, in one case, 30 points in the minority. On the whole, he believed in the constancy of the quotient. He stated that the practice in Northumberland was now to admit children to Secondary Schools who were very good either in intelligence tests or in acquired knowledge, provided they were at least average in the other respects. At the free place examinations in 1921, however, 14 children were awarded free places and maintenance on the result of the intelligence test alone. Of those one child had since died and one was doing badly, but the remainder were doing very good work in Secondary Schools. His experience in Northumberland had led him to reject for general educational use group tests which required accurate timing for each page, partly because the bustling nature of such tests was not in his opinion desirable, and in part because considerable skill and coolness were required in the supervisory, so that testing on a large scale (in which it was necessary to employ untrained supervisors) was impossible with such minute timing[1].

(*b*) In a memorandum sent by the Director of Education for Northumberland it is stated that, in addition to the tests applied in 1921 and 1922, the Education Committee has sanctioned the experimental adoption of intelligence tests in examinations for Junior Scholarships, and as a supplement to attainment tests for the admission of fee-paying pupils to Secondary Schools. For this purpose the following group tests have been used :—

Year.	*Examination.*	*Test Used.*	*No. of Candidates.*
1921	Admission to Secondary School.	Prepared by Mr. C. A. Richardson.	472
May, 1922	Award of Junior Scholarships.	Tests prepared by Dr. Godfrey Thomson.	1,714
1922	Admission to Secondary Schools.	Terman Group Test, Form A	392
1923	Award of Junior Scholarships.	National Intelligence Test, Scale A.	1,900
1923	Admission to Secondary Schools.	National Intelligence Test, Scale B.	443
1923	Intending Teachers ..	"Simplex" tests, prepared by Mr. C. A. Richardson.	218

[1] Cf. Professor Thomson's description of this Northumberland experiment in the *British Journal of Psychology* for December, 1921, and July, 1922.

As a general rule, candidates who obtained satisfactory marks in the subjects of English and Arithmetic proved also to have shown, as the result of the intelligence test, a mental ratio above the average. To this rule, however, there were many instructive exceptions.

The Director of Education has submitted a number of tables showing the result of an inquiry into the subsequent progress of pupils admitted to Secondary Schools during the past two or three years. The following are the more significant facts emerging :—

(1) Among pupils credited by the intelligence tests with a mental ratio of 120 or more, all except one are reported as making satisfactory progress. In the ordinary examination in English and arithmetic many of these brighter candidates obtained low marks, and would not have received a scholarship had not the results of the psychological test justified special consideration. They are all well under the average age of the forms in which they are now working; and their position in their form is high—some taking first, second or third place in spite of low marks in the ordinary examination. One boy who obtained ratio of nearly 120, but only 15 marks out of 100 in the ordinary examination, is now second in his form, although he is a year younger than the rest of the class : the head master reports that " though he failed in paper work on entrance, he has since made good."

(2) Three pupils who showed somewhat low mental ratios (between 83 and 90) were awarded scholarships on the ground of comparatively high marks in arithmetic and English. One is reported to be " a steady worker " ; and the work of the other two is described as " fair " or " very fair." They have, however, obtained only low places in their forms ; and two of them are much above the average age of their form. A fourth boy, with an equally low mental ratio, whose examination marks did not justify a scholarship, is reported as " satisfactory."

There are thus exceptions to the rule that only children with high mental ratios make satisfactory progress in a Secondary School, but these exceptions appear to be exceedingly rare.

The Director concludes that " on a broad view, the introduction of a psychological test, as a supplement to the examination in school subjects, has been of great value. It has at least tempered the application of a particular method of determining admission to Secondary Schools with entirely satisfactory results."

(5) *An Experiment in the use of Group Tests of " Intelligence " in* 16 *Elementary Schools in the West Riding of Yorkshire* (*summarised from a memorandum sent by Mr. T. P. Tomlinson, M.Ed., M.A., Head Master of Love Lane Council School, Pontefract.*

Group tests of intelligence were applied to 1,600 children, aged 9 to 14, attending 16 schools in the West Riding of Yorkshire. The

area thus surveyed included ten small villages, a small river port, and slum and residential districts of a neighbouring city.

The scheme of tests used was one termed by Mr. Tomlinson the West Riding Scale. It consisted of eight tests of familiar types :— (1) Instructions, (2) Analogies, (3) Absurdities, (4) Mixed Sentences, (5) Classification, (6) Word-meaning, (7) Logical Selection, and (8) Arithmetical Reasoning. Two parallel sets of similar tests were used.

The results of the experiment are briefly as follows :—

(1) It was found that all the tests could be administered satisfactorily with a minimum of explanation by the head teachers of the various schools. The application of the tests took rather less than one hour ; and the marking was practically mechanical.

(2) This is one of the few experiments in which similar tests have been twice applied, so that some measure of the correspondence of the results is obtainable. The results of the second application agreed almost exactly with the results of the first, the reliability coefficient being throughout approximately ·95.

(3) For a few small age groups independent estimates were obtained, either from the teachers of the classes, or by the application of the individual tests of the Binet-Simon Scale. The correspondence of these estimates with the group tests was again quite close, the correlations ranging from ·71 to ·84.

(4) A point of special interest in the survey is the well-marked difference found between children from areas of different types. The results obtained from the rural schools are of peculiar interest. In the rural schools the most noticeable feature is the greater range in the distribution of intelligence. Further inquiry showed that the villages where a disproportionate amount of dull and backward children were found, were for the most part villages near a port or city—villages, that is to say, from which the brighter families had already migrated. Other villages, however, in a similar situation showed a disproportionate number of bright and average children : in such cases it was found that the less intelligent families had moved to the larger industrial centres owing to the demand in the latter for unskilled labourers ; while, on the other hand, the village itself offered a desirable place of residence for better families, such as skilled artisans and professional people.

(5) Mr. Tomlinson points out that such tests, in the hand of teachers working in relatively isolated schools, afford serviceable aid in rapidly sorting out the dull, the backward, and deficient, from the bright and super-normal. Without such standardised tests teachers in these remoter districts have little notion of the general level of the normal, backward, or super-normal child. He urges that such tests are needed more particularly for the brighter children. The sub-normal are readily detected owing to their

backwardness in school progress, but the brighter children may easily be overlooked.

(6) The following are typical instances of children whose intelligence and fitness for higher education have been wrongly rated by the ordinary examination.

(i) A.B., a girl of 11½, had a mental ratio of 147 (Binet) and 148 (West Riding Tests). Her head mistress described her as a brilliant girl. Yet she had failed in the Junior Scholarship Examination. She was subsequently transferred to a Secondary School, and is now working there very successfully.

(ii) C.D. had a mental ratio of 125, but is a highly nervous child. At the age of 6 he was able to read the daily newspaper with intelligence. At 8 his criticial ability was so far developed that he was able to suggest improvements in the designs exhibited in his "Meccano" book, and to embody these improvements in his constructions. At 11 he was doing very good work in Standard VI, both in arithmetic and English. Owing to his nervousness, however, he failed hopelessly in arithmetic at the Junior Scholarship Examination.

(iii) E.F. has gained a Junior Scholarship, owing, probably, to special preparation; his low mental ratio should have eliminated him from the list of awards. His most recent report shows that at the Secondary School he is bottom but five in the lowest form, and his want of success, even in subjects taught in the Elementary Schools, is almost as evident as it is in the new subjects taught at the Secondary School.

(7) In a central school included in the survey, nearly 2 per cent. of the children had mental ratios below 85. Owing probably to special preparation, these children had been selected for a central school on the basis of the ordinary examination in scholastic attainments. Though at the time of the examination by no means backward in school work, they are yet, as the tests show,distinctly dull. The head master's subsequent report on these children shows that they are unable to make normal progress in the central school.

(6) *Results Obtained by the Use of Individual and Group "Intelligence" Tests at the Ben Jonson Council School, Harford Street, Stepney, London (summarised from information supplied by the Head Mistress of the Girls' Department, Miss G. M. Anderson).*

Miss Anderson explained that this school was the largest provided Elementary School in London. Eighty per cent. of the pupils were the children of casual workers, and about 12 per cent. were of foreign parentage.

For the purpose of classifying scholars on admission, individual tests were generally used—Terman's revision of the Binet-Simon Scale. The test was applied during first term to pupils admitted from the Infants' Department, and immediately on admission to those admitted direct from outside.

Group tests were also used such as Dr. Ballard's " picture," " absurdity " and " vocabulary " tests, but the individual tests were considered the best. Of these individual tests, the " comprehension " tests were the most reliable.

The school was organised so as to allow three currents of promotion, viz., the " super-normal," with a mental ratio of 110, the " normal " with a mental ratio of 90 to 110, and the " sub-normal " with a mental ratio below 90. Mental tests formed the sole basis of classification for the super-normal grade. No test was required in order to discover the sub-normal children.

Promotions in the school were determined as a result of examinations held twice a year. The examination was chiefly in acquired knowledge, but mental tests in the form of " comprehension " tests were always included. The order of merit in the two tests was on the whole parallel, especially in the super-normal class. The following table was given as a typical result of an examination in the Upper Division of the " Super-normal " class :—

Position on Class Register.	*Intelligence Quotient.*	*Remarks.*
1	127	Gained a Junior County Scholarship.
2	131	Do.
3	116	Do.
4	116	Gained a Foundation Scholarship.
5	120	Gained a Junior County Scholarship.
6	117	Gained a Foundation Scholarship.
7	115	Do.
8	110	Transferred to " normal " class.
9*	130*	To be transferred to a Central School.
10	118	Do.
11	120	Do.
12	113	Do.
13	122	Gained a Junior County Scholarship.
14	112	To be transferred to a Central School.

The time taken to work the tests was regarded as indicative of intelligence, but the slow, sure, deliberate and accurate children were also discovered, and one such child had proved to be brilliant.

* This child was rather inaccurate in arithmetic.

(7) *Results Obtained by the Use of Group Tests of Intelligence at the Mile End Central School, Myrdle Street, London (summarised from data supplied by Mr. J. G. Tibbey, M.A., Head Master of the School and Chairman of the Research Committee of the London Head Teachers' Association).*

Mr. Tibbey had used psychological tests in two boys' schools and one central mixed school, sometimes to assist in organisation, and sometimes to test it. For this purpose he had mainly relied upon Opposites, Analogies, Arithmetical Series, and Absurdities, and, allowing for errors in application, had found such tests materially helpful. Much depended upon care in the selection of tests and attention to procedure.

For the purpose of school grading at the present stage, it was advisable to consider the results obtained by intelligence tests in conjunction with those of the examinations in attainments and the class record.

Time tests were generally valuable, but needed to be applied with discretion. They appeared to be undesirable for pupils below 9 years of age where writing was involved.

He suggested that copies of intelligence tests, of proved reliability should be made available to teachers by local education authorities.

(8) *Account of an Experiment in the Use of Individual Non-linguistic Tests of Intelligence and Quasi-vocational Tests at the Acton and Chiswick Junior Technical School (summarised from information supplied by the Head Master, Mr. Max Tagg, B.Sc.).*

During the past four years, Mr. Tagg has regularly applied psychological tests to all the boys in his school. The boys are about 150 in number, and range in age between 13 and 16 years. The tests used were designed to test partly general intelligence and partly special aptitude for engineering trades. They consisted of individual tests of a non-linguistic or so-called "performance" type, e.g., tests of perception of form and of space, of gauging curves, diameters, lengths of lines and sizes of angles, of constructing geometrical figures, and of practical reasoning.

After considerable experience of these tests, he concluded that they afforded better estimates than the ordinary written examination. In a written examination boys of the type attending his school did not always do themselves justice, and sometimes failed to reveal their special aptitude for technical work. Often there was little or no correlation between the results of the written examination and the opinion which masters formed of individual boys during their attendance at the school. Mr. Tagg himself now bases his final judgment of his pupils on the combined result of written, oral, and practical examinations, in all of which psychological tests play a prominent part.

(9) *An Experiment in the Use of Tests at Pate's Grammar School for Boys at Cheltenham (summarised from information supplied by the Head Master, Mr. R. R. Dobson, M.A.).*

(1) *Oral tests of intelligence.*—Mr. R. R. Dobson carried out in 1922 an experiment in supplementing the ordinary annual examination for free places by an oral examination with "intelligence" tests. There were about 150 candidates for 26 places. The tests used for this purpose were Dr. Burt's graded reasoning tests.

From his personal experience in using these tests he concluded that they were undoubtedly of great value as a check upon the scholarship examination and as a useful subsidiary criterion for borderline cases.

In the main, the agreement between the psychological tests and the ordinary examination was so close that, in many Elementary Schools, the boys could have been selected by the results of the former alone. Since, however, a child must have some degree of educational attainments, it would clearly be undesirable to abolish the ordinary examination.

Where the results of the psychological tests differed from those of the ordinary examination he had inquired into the causes, and found cases of two types :—

(*a*) The teaching in the Elementary Schools in a mixed area varied far more than was generally supposed, some being very efficient and some inefficient. Thus there could be little doubt that many children who did better in the psychological tests would have done equally well in the ordinary examination had they been as efficiently taught as the rest. In certain country schools, for example, the boys were, in natural intelligence, quite as well endowed as the town children ; but, owing to the difference in staffing, often attained a far lower standard in the ordinary examination. Mr. Dobson had accepted some of these boys solely on the ground of their success in the "intelligence" tests, and their subsequent progress showed that the selection had been fully justified.

(*b*) Where a boy from a good home or school had failed to do the work of the class to which, owing to his more attractive bearing, or his greater success in the scholarship examination, he had been sent, such tests had been of great service in indicating that his apparent merits were simply a superficial and transitory effect of fortunate conditions, and in distinguishing, in such instances, lack of native ability from lack of industry and zeal.

(2) *Written tests of intelligence.*—Mr. Dobson has since carried out extensive experiments with written or group tests of intelligence. The tests employed consisted of five separate tests, constructed by Dr. Burt, practically identical with those used later by Mr. Vaughan at Rugby School. With the written tests of intelligence

he has examined nearly 600 persons, including boys and girls from Pate's Grammar Schools at Cheltenham, students in the training colleges and University at Bristol, and a small group of university lecturers and professors.

The written tests gave a correlation with attainments much closer than that given by the individual tests, the correlation being somewhat higher with the younger grammar school pupils than with the older university students. Mr. Dobson believed that the results of the oral and individual tests were the more reliable as indications of native intelligence ; but the written tests, which took far less time, might for practical purposes be more serviceable.

The following is an extract from the tables submitted by Mr. Dobson to show the agreement between the results of the tests and the pupils' progress in the ordinary work at the school.

Boys : *Ages* 11–12.

Intelligence.	*Place in Class.*
1	1
2	2
3	3
29	33
31	29
32	30
33	35
34	34
37	36
38	38
39	39
42	42
43	40
44	37

Among the group aged 15 to 16, the three top boys in the test were all matriculated students, doing well in an advanced course. K.L., aged 11, who obtained 132 marks (50 marks above the average for his year, and by far the highest obtained among boys of his own age), is known as an unusually bright boy, working well with grammar school boys whose age is 14. M.A.P., aged 17, who gained the high mark of 166, almost the highest in the whole school, had obtained a State Scholarship at the age of 15, with distinction in mathematics in the Higher Schools Certificate, and has since obtained an open mathematical scholarship at Clare College, Cambridge.

Mr. Dobson found that inborn intelligence, as thus tested, appeared to increase regularly from the lowest ages tested up to about the age of 16 ; after that age there seemed little or no increase in intelligence as distinct from experience and attainment, except among persons of exceptionally high ability ; the university lecturers

and professors, for example, obtained unusually high marks, well above those obtained by fourth-year students in training who had already obtained their degree with honours.

Mr. Dobson stated that the boys admitted into the school in 1922 graded by the oral reasoning tests are all doing good work in the second year of the school course.

At the re-examination of these boys at the end of one year's interval, all kept their places well. There appeared to be a slight increase in intelligence quotient, possibly due to scholastic training and environment.

(10) *Memorandum on an Experiment in the Use of Group Tests of "Intelligence" at Rugby School by the Head Master, Mr. W. W. Vaughan, M.V.O., M.A.*

"On the last day of the Christmas Term, 1923, a group test was set to the whole of Rugby School (600 boys). The test done was one constructed by Dr. Burt, for the National Institute of Industrial Psychology, and consisted of series 33 of their group tests. The series consisted of a printed booklet of five different tests—namely, Opposites, Analogies, Mixed Sentences, Completion, and Reasoning—similar to those described in Appendix VII. To make the marking as uniform and as mechanical as possible, alternative answers were printed, so that the candidate had nothing to write, but simply to pick out and underline the correct reply. Such a test necessarily affords no opportunity for constructive imagination or literary expression, but was selected as being the type of test most suitable for testing a large group of boys differing widely in age and ability. In a previous year a similar test had been employed on a smaller group of boys at Rugby School, and the results were found to correlate with the masters' independent estimates of intelligence to the high extent of ·83. The work was done Form by Form, under very strict conditions, and I think the instructions were most loyally obeyed. The scripts were sent to be corrected by a clerical assistant nominated by Dr. Burt. The results are interesting. From the tables given it will be seen that the upper forms of the school did better than the lower forms; the older boys better than the younger boys; those who hold scholarships better than those who are normally considered backward in their work. Of those who came out in the first 20 all except two were in the sixth form, or the adjoining form. The school position of one boy who was in the last four of those examined was about 150th from the bottom by school work, but he has gained this place by very hard work. He finds all his subjects difficult except Greek Grammar, in which he does very well.

"From my knowledge of all the boys in the school I conclude that the test is a very good test of teachable ability. This seems to me a better word than "intelligence." In a few, but very few, cases, the boys were upset by the time limit.

"The test, it is true, gave no reliable indication of a boy's power of literary expression, or of his imagination. It did, however, give a better judgment than I had expected of his industry. As a rule a boy who does not waste time in school or in preparation was more successful than the dawdler. The results seem to disprove the contention that there is very little change in a boy's intelligence between 16 and 18.

"It may be worth while to add that amongst the members of the school who took the tests were a Scholar of Balliol, a Scholar of Trinity College, Cambridge, and a Scholar of King's College, Cambridge."

TABLE A.

Forms.

Upper School.

Maximum Mark = 200.

Classical Side.			*Modern side.*		
	Average Age.	Average Marks.		Average Age.	Average Marks
VIa	17 yrs. 8 mo. ..	$163\frac{1}{2}$	VI	17 yrs. 8 mo. ..	$160\frac{1}{2}$
VIB	17 yrs. 11 mo. ..	$159\frac{7}{8}$	XX	17 yrs. 4 mo. ..	$149\frac{7}{8}$
XX	16 yrs. 4 mo. ..	$156\frac{5}{8}$	VI	17 yrs. 1 mo. ..	$137\frac{1}{2}$
Va	16 yrs. 2 mo. ..	$141\frac{5}{8}$	LVB	16 yrs. 6 mo. ..	$145\frac{5}{8}$
LVa	15 yrs. 6 mo. ..	$154\frac{7}{8}$	LVc	16 yrs. 5 mo. ..	$136\frac{1}{13}$

Army Class, 17 yrs. 5 mo. .. $137\frac{1}{2}$

Middle School.

	Upper Middles.		*Lower Middles.*	
I	A.	15 yrs. 9 mo. $140\frac{7}{8}$	A.	15 yrs. 3 mo. $114\frac{1}{4}$
	B.	16 yrs. 0 mo. $143\frac{7}{8}$	B.	14 yrs. 11 mo. $120\frac{1}{2}$
II	A.	15 yrs. 8 mo. $128\frac{1}{2}$	A.	14 yrs. 8 mo. $119\frac{5}{8}$
	B.	15 yrs. 4 mo. $131\frac{1}{2}$	B.	14 yrs. 5 mo. $121\frac{7}{8}$
III	A.	15 yrs 1 mo. $129\frac{7}{8}$	A.	14 yrs. 5 mo. $114\frac{1}{2}$
	B.	14 yrs 11 mo. $132\frac{1}{2}$	B.	14 yrs. 6 mo. $109\frac{1}{2}$

Lower School, 13 yrs. 11 mo. .. $105\frac{3}{4}$

TABLE B.

Average mark for the whole school	134½
Average mark gained by 20 oldest boys in the school ..	150¼
Average mark gained by the 20 youngest	112½
Average mark gained by 40 scholars	159
Average mark gained by 40 liable to superannuation ..	127
Average mark gained by 20 best musicians	144
Average mark gained by 20 best scholars (classical) ..	163
Average mark gained by 20 best mathematicians	163
Average mark gained by 20 best science boys	161
Highest mark gained by any boy	182
Lowest mark gained by any boy	66

APPENDIX III.

Notes by the Secretary on the Use of Psychological Tests of various Types in Foreign Countries.

Austria.

(A) *Tests of Intelligence.*

It is reported that in 1919 admissions to "High Schools" (i.e. Universities and Technical and Commercial Colleges) were made largely on the basis of psychological tests.

An order of the Ministry of Education dated 15th May, 1922, made it permissible for Local Education Authorities to use in Elementary and Higher Elementary Schools a form of description of each individual pupil which was to serve as a guide in the selection of a higher type of school or a trade. This "Schülerbeschreibung" was intended to give detailed information regarding the physical characteristics of the pupil, his progress in the various school studies and his mental and psychical characteristics. It was to be filled in by the teacher and used experimentally during the school years 1922–23 and 1923–24[1].

(B) *Vocational Guidance.*

The Ministry of Social Administration at Vienna, established on 1st January, 1918, undertakes to give vocational guidance to children leaving school. Important researches and experiments in the use of vocational tests are conducted in the psychological laboratories attached to the Universities of Vienna and Innsbrück[2].

---

[1] *Volkserziehung* for 1st September, 1922, Stück, xvii.

[2] Fontègne. *L'Orientation professionelle*, Paris, 1923, pp. 136–137. International Labour Review for May 1922, Vol. V, p. 713.

Belgium.

(A) *Tests of Intelligence.*

Tests of intelligence are used experimentally in a few elementary and secondary schools. They are employed in Brussels by Dr. Demoor as an adjunct to the examination of mentally defective children[1].

(B) *Vocational Guidance.*

The Intercommunal Bureau for the vocational guidance and apprenticeship of the youths and girls of Greater Brussels was originally opened in July, 1914, on the initiative of the Société Belge de Pédotechnie. It was re-opened after the war in December, 1919. It consists of a library, a laboratory and three branches, medical, psychological, and technical, each under a director. The director of the technical branch is also general director of the bureau which is supported by the several communes of Greater Brussels, each of which has agreed to contribute for five years the income accruing from an annual rate of two centimes per inhabitant. The different communes of Greater Brussels are now establishing their own bureaux for educational guidance equipped with the services of medical, psychological, and educational experts. These work under the general supervision of the Intercommunal Bureau, the functions of which include the instruction of the communal bureaux in the principles of vocational guidance, the publication of tabular statements showing the various forms of ability required for different occupations, and the examination of special cases at the request of any of the communal bureaux. The general system of vocational guidance is based on an elaborate medical examination, a full psychological investigation, and a detailed report to a questionnaire from the teacher. A systematic " follow up " is carried out in order to test the accuracy of the tests employed and to enable modifications to be introduced where necessary. The Bureau is assisted by a Committee[2].

China.

The National Association for Vocational Education at Shanghai, was founded " to establish proper relations between education and vocation." The Association conducts research into educational and industrial questions and encourages vocational guidance with the aid of psychological tests. Several Vocational Schools have recently been established in China[3].

---

[1] From information supplied by Monsieur A. G. Christiaens, cf. J. Demoor et T. Jonkheere, *Science de l'education*, Brussels, Lamartin, 1922.

[2] *Bulletin Trimestriel de l'Office Intercommunal, etc.*, Brussels, Lamartin, 1923, cf. also *Journal of the National Institute of Industrial Psychology*, Vol. I, p. 31 and p. 212.

[3] From information supplied by Dr. G. H. Miles.

Czechoslovakia.

(A) *Tests of Intelligence.*

The Binet-Simon Scale was translated into Czech with slight modifications in 1911, and has been extensively used in the Bohemian Paedological Institute (Český Pedologický Ústav) founded at Prague in 1910. The Institute is now organised in four Sections, as follows:—

(*a*) The Pædometric Section, which takes elaborate medical measurements, etc.
(*b*) The Educational Section.
(*c*) The Psychological Section.
(*d*) The Pathological Section (for examining and treating children who are mentally or physically defective).

The psychological tests in use in the Psychological Section include a specially constructed set of ten tests of intelligence, which are applied every year to about 700 children of 11 years of age.

Similar Institutes have recently been established at Brno (Brünn) and Liberec (Reichenberg)[1].

(B) *Vocational Guidance.*

Centres for vocational guidance were established from 1920 at Prague, Brno (Brünn), and five other towns.

It is reported that in 1920 over 1,000 children were examined in these centres partly by means of psychological tests.

The headquarters are at Prague, and the work comprises research into mental and physical capacities and abilities, and provides for examination of the capacities required for various branches of industry and for the army. The Prague Office is divided into a Vocational Section and a Students' Section. Two thousand and seventy-eight young persons were tested in 1923[2].

The Institute for Applied Psychology at Prague, organised in 1921 as a department of the Masaryk Academy of Labour, is conducting researches in regard to vocational tests and the requirements of various industries[3].

Denmark.

Vocational Guidance.

In connection with the Municipal Labour Exchange in Copenhagen, a Laboratory for Vocational Testing is being organised for boys and girls. The expenses are met by the municipalities of Copenhagen and Friederiksberg, aided by a State Grant. The Laboratory will be opened on June 1st, 1924.

---

[1] Article by Dr. Kyril Stejskal in *Zeit. für pädagogische Psychologie*, 23 Jahrgang. Heft. 11–12 November–December, 1922, pp. 453–458.

[2] From information supplied by Dr. Mauer.

[3] From information supplied by Prof. Forster. Accounts of the work of this Institute are published in *Nova Prace*, the organ of the Academy of Labour. Cf. also *Ceska Technika* for June, 1921.

The Danish Chamber of Manufacturers and the Employers' Federation are considering the question of establishing a Laboratory for vocational testing in association with the Technical College at Copenhagen[1].

Finland.

A psychological laboratory for vocational guidance has been established at Helsingfors, primarily for the selection of apprentices for the State railways. The work is maintained by the Department of Professional Teaching for State Railways[2].

France.

(A) *Tests of Intelligence.*

Though Binet and Simon were Frenchmen their celebrated Scale does not appear to be very extensively applied in France outside of Paris and some of the larger towns, where it is chiefly used for segregating defective children for special treatment[3].

(B) *Vocational Tests.*

At Bordeaux there is a private organisation, the Chambre des Métiers, which carries out psychological work in vocational guidance[4].

At Nantes l'Office Regional de la main d'oeuvre gives vocational guidance and assists in the placing of apprentices. It is supported by grants from the Ministry of Labour, the City of Nantes, the local Chamber of Commerce and the Technical Educational Authorities. The Office state in a recent report that the main conclusions arrived at as a result of their application of psychological tests for purposes of vocational guidance are as follows :—

(*a*) That vocational guidance is essentially a work of education, which, in order to produce its full results ought to be applied to the pupil many years before he leaves school.

(*b*) That school records are essential and should be developed along definitely psychological lines.

(*c*) That success in vocational guidance depends on the due co-ordination of the efforts of teachers, psychologists, doctors, and employers.

---

[1] From information supplied by Mr. Kai Mygind.

[2] From information supplied by Dr. G. H. Miles.

[3] Cf. *L'Année psychologique*, passim.

It is worth noting that a proposal was made in an Education Bill presented (by a private member) to the Chamber of Deputies in autumn, 1920, that the Binet tests should be employed, in conjunction with the ordinary school examinations and the school records (*Livrets Scolaires*) in selecting Elementary School children of 12 for further education whether Secondary or Technical. The Bill was not passed. Cf. *Le Temps* of 20th October, 1920, p.1.

[4] J. Fontègne, *L'Orientation professionelle*, p. 144, cf. also *La Rose des métiers*, Paris, 1922, by F. Mauvezin, the Director of the Bordeaux Chambre des métiers.

(*d*) That the whole movement for vocational guidance should be centralised and a clearing house for information established[1].

Similar organisations exist at Lyons, Marseilles and Roubaix.

At Paris the Institut Lannelongue conducts *inter alia* researches bearing on the problems of vocational psychology.

At Strasbourg there is a special Bureau for vocational guidance for Alsace and Lorraine under the direction of Professor J. Fontègne[2].

The idea of using psychological tests as an aid to vocational guidance appears to be gaining ground in France. Professor J. Fontègne advocates the use of such tests in the various "Offices régionales d'orientation professionnelle," which are now being organised throughout France, with the encouragement of the Technical Department of the Ministry of Public Instruction[3].

Germany.

(A) *Tests of Intelligence and Standardised Scholastic Tests.*

Tests of intelligence and standardised scholastic tests are extensively used in various towns, e.g. Berlin, Charlottenburg, Hamburg, Altona, Göttingen[4], Hanover[5], Breslau, Guben in Brandenburg[6], Freiburg in Breisgau, Erfurt, Dresden and Schneeberg in Saxony[7], in connection with the selection of children from elementary schools for admission to secondary, quasi-secondary, and technical schools. In Hamburg, for example, over 1,000 children of ten years of age have been examined every year in the spring since 1918 by means of intelligence tests with a view to transfer to certain higher classes and schools[8]. At Altona the tests have been used to assist in the classification of children of about seven years of age in elementary schools[9]. In Berlin performance and other tests have been em-

---

[1] From information supplied by Dr. G. H. Miles.

[2] Fontègne, op. cit., p. 6. The scheme of the psychological examination in use at this Bureau is printed in Fontègne, op. cit., pp. 235–236.

[3] Fontègne, op. cit., pp. 58–75, cf. also E. Gaultier's article on Vocational Guidance in *International Labour Review*, Vol. V, No. 5. (May, 1922), p. 714.

[4] R. Peter und W. Stern ; *die Auslese befähigter Volksschüler in Hamburg*, Leipzig, A. Barth, 1922, *passim.*

[5] Cf. article by Dr. W. Hische, die Auslese der Begabten in Hanover in *Praktische Psychologie* for February, 1921, pp. 129–142.

[6] Cf. article by G. Hirsch in *Praktische Psychologie* for June, 1922, pp. 268–278.

[7] Article by A. Schwärig, Zwei Begabungs-prüfungen. in *Zeit. für Pädogogische Psychologie*, pp. 430–441.

[8] T. Fontègne, *L'Orientation professionelle*, p. 130, Paris, 1923.
Peter und Stein, op. cit., pp. 157–161 and *passim.*, cf. also *British Journal of Psychology*, Vol. XI, Part 2 for January, 1921, pp. 251–253.

[9] W. Stern, *Das psychologische Laboratorium der Hamburgischen Universität*, Quelle und Meyer, Leipzig, 1922, pp. 33–34.

ployed to select pupils of 13 to 19 years of age with special gifts for artistic design[1].

At Berlin Dr. Curt Piorkowski, Director of the Orga Institute, acting at the special request of the Civic Educational authorities has for some years past applied psychological tests to the candidates for entrance to the two Secondary Schools for exceptionally gifted boys, (Begabtenschulen) established in 1917 by the City of Berlin. There is very keen competition for entrance to these schools. Dr. Piorkowski, in association with Professor Moede, the Director of the Psychotechnical Laboratory at Charlottenburg, applied certain tests to all candidates with a view to selecting the most suitable. The fundamental requirements expected of all candidates were adequate power of concentration, good memory for logical construction, ability to give clear and unambiguous decisions regarding given situations, and creative imagination. The psychological tests employed were specially designed to discover the existence of these four capacities, and were expressly devised with the object of determining natural abilities, not knowledge acquired at school. It is stated that the candidates selected by this method have already displayed undoubted signs of high promise. Nevertheless, it is evident that definite inferences cannot be drawn until the lapse (say) of ten years, when a review of the progress made by these boys will be possible[2].

Similar tests have been applied to candidates for the Berlin Begabtenschulen for girls.

" Intelligence " tests have also been used experimentally in the examination for admission to a Training College for Women Teachers at Hamburg[3].

(B) *Vocational Guidance.*

Brief accounts of some typical Institutes and other organisations for research on psychological tests with special reference to vocational guidance.

(i) The Psychological laboratories attached to several of the Universities, e.g. Hamburg[4], are conducting important experiments

[1] Dannenburg's article on Auslese and Berufsberatung der künstlerisch Begabten in *Praktische Psychologie*, No. 5 (1920), pp. 150–159.

[2] Moede, Piorkowski and Wolff, *die Berliner Begabtenschulen*, Beyer, Langensalza, 1919. *Zwei Jahre Berliner Begabtenschulen, Erfahrungen ihrer Schulleiter*, Hirzel, Leipzig, 1920. Cf. also article by Dr. Piorkowski in *Journal of National Institute of Industrial Psychology*, Vol. I., No. 3. (1922), pp. 99, foll. and article by Margaret Steppat in *Journal of Experimental Pedagogy*, Vol. 6, Nr. 3. (1921).

[3] Article by A. Penkert, über die Anwendung von Tests bei Anfnahmeprüfungen in ein Hamburger Lehrerinnenseminar published in *Hamburger Arbeiten zur Begabungsforschungen*, Nr. 11, Leipzig, 1919.

[4] See W. Stern, *Das psychologische Laboratorium der Hamburgischen Universität*, Leipzig, 1922, pp. 25 foll.

in the construction and application of vocational tests and tests of intelligence.

(ii) *The Psychological Laboratory of the Vocational Institute at the Technical College, Stuttgart.* The Psychological laboratory attached to the Vocational Institute at the Technical College at Stuttgart aims at testing children and adults for occupational fitness. It also tests candidates seeking admission to the Technical Colleges. It is associated with the Psychological laboratory attached to the local Technical College and is maintained by the State of Würtemberg and the City of Stuttgart[1].

(iii) *The Psychotechnical Institute at Charlottenburg.* The Psychotechnical Institute attached to the Charlottenburg Technical College collaborates with a number of great industrial undertakings such as the Osram, helping them to select and classify their pupils and operators.

The fullest use is made of information such as school records, medical data and so forth. Inquiries are also made regarding the candidate's interests and leisure occupations. Every endeavour is made to secure the co-operation of each applicant in the work of determining his suitability for a future career.

Group tests consisting of modifications of well-known " intelligence " tests, memory for number, form, and so forth are applied together with tests designed to indicate technical ability. For example, the pupil is shown a drawing and a simple arrangement in levers and his ability to foretell what would happen on moving one of the levers is determined. His mechanical ability is gauged by means of specially designed apparatus which has been standardised and patented. Steadiness of hand is measured by various tests, for example, his ability to pass a thin rod through holes without touching the sides. As the result of the application of a number of such tests a considerable amount of information regarding a pupil's abilities can be obtained, from which a useful forecast of his suitability for engineering work can be prepared.

The candidate's subsequent career is closely followed, and any discrepancies between the results obtained in the psychological laboratory and his actual achievements in school and workshop are most exhaustively examined, the information thus obtained being used to correct future placings, or being made a subject of further research. It is stated that this procedure has the full support and sympathy of the local Trade Unions, which realise that it is to their advantage to have within their organisation only men who are really

---

[1] From information supplied by Dr. G. H. Miles.

capable. A contribution towards the support of the Institute is voted annually by the Unions, and the satisfactory passing of these tests has been made a condition of apprenticeship.

Similar psychotechnical laboratories have been established at Leipzig, Munich, Frankfort on the Main and other large cities[1].

(iv) *The Orga Institute at Berlin.* The work of this Institute, recently established in Berlin includes vocational guidance and selection.

The Institute possesses a laboratory, well equipped with apparatus for vocational testing[2].

(v) *The Berlin Institute for Applied Psychology.* This Institute, founded in 1906, was amalgamated in 1918 with the Institute of Industrial Psychology, founded in 1916, and is now under the direction of Dr. Otto Lipmann.

It serves as a central clearing-house for information regarding the different applications of psychology and possesses a laboratory for research work on psychological tests[3]. It publishes a periodical *Zeitschrift für angewandte Psychologie.*

(vi) *The Institute of Practical Psychology at Halle.* This Institute began by giving vocational guidance to disabled soldiers, but now also gives such guidance to normal and abnormal children. Special stress is laid on continued psychological observation of children during their school course[3].

(vii) *Berufsamt der Stadt Berlin* (*Vocational Bureau of the City of Berlin*). In addition to a Statistical Department and an Employment Exchange, the Bureau has a Psychological and Medical Department where information concerning each candidate's physical and mental capacities is collected, psychotechnical tests are applied, and medical examinations are carried out.

The Bureau issues leaflets briefly describing occupational conditions and appeals to children and their parents to ask advice when making a choice of vocation in life. It emphasises the need for applying for work and for improving educational attainments. A choice of vocation should be made according to the child's liking for the work, and according to health, physical development, and special aptitudes. The Bureau also suggest that its officials should discuss the matter

---

[1] *Journal of the National Institute of Industrial Psychology,* Vol. I, p. 16 and pp. 190–192.

[2] *Journal of the National Institute of Industrial Psychology,* Vol. I, p. 110.

[3] From information supplied by Dr. Otto Lipmann.

with parents, taking into account school records and other information about each individual child. The Psychological Section of the Bureau conducts its work as follows :—

The applicant has generally decided on the type of work he desires, and the aim of the Bureau is to ascertain the nature of his choice. For this purpose his parents are interviewed, the opinions of his teachers are obtained, and the records of work at school and any outside work accomplished by him are considered. The Bureau state that a critical examination of school records showed that 30 per cent. were of very little use and 10 per cent. quite useless. This discovery led to the formation of a scheme for keeping definite observation records of each candidate, which were proving to be of real value.

The physical history of each candidate is ascertained from the reports of the School Medical Officer.

Psychological tests of different types are then applied to the candidate on two days, consisting of :—

(*a*) Group " intelligence " tests to ascertain the level of intelligence together with tests devised to discover occupational wish.

(*b*) These tests are followed up by individual " intelligence " tests.

The " psychological profile " (psychogram) of each applicant is obtained from the data yielded by the group tests and individual tests. The experts who have tested the candidates then discuss the results obtained from (*a*) the " psychological profile " and (*b*) their own observation records. The results are set out on a standard chart, which is forwarded to the Vocational Guidance Department for final use. It is stated that many large firms are interested in the work of the Bureau and afford opportunities for following up the results[1].

There are similar municipal organisations at Bonn, Dresden, Leipzig and other towns.

(viii) Several great industrial concerns, e.g. the Krupp Company at Essen and the A.E.G., have psychotechnical departments attached to their works in which vocational tests are applied to the operatives. In the same way the Imperial Railways Administration and the Berlin Tramways apply psychological tests to their employees[2]. Similar tests are applied to apprentices in the printing trade at Munich[3].

---

[1] From information supplied by Dr. G. H. Miles.

[2] *Praktische Psychologie*, Jahrgang I, pp. 309, foll. and *Zeit. für angewandte Psychologie*, Band 18, pp. 107 foll.

[3] A. Huth, *Die Münchener Eignungsprüfung für Buchdrücker*, Leipzig, 1922.

Holland.

(A) *Tests of Intelligence and Standardised Scholastic Tests.*

Tests of intelligence are applied to pupils entering the Amsterdam Lyceum and several other large Secondary Schools in Holland. Experiments in the use of tests are carried out by Dr. Brugmans at the Dr. D. Bos Foundation of the Pedagogical Institute at Groningen[1].

(B) *Vocational Guidance.*

The Municipal Bureau for choice of vocations at Amsterdam, founded in 1920, applies intelligence tests to some of the youths and girls seeking vocational guidance.

The Psychotechnical laboratory at Amsterdam, opened in 1921, is a municipal institution connected with the department of Public Health. Among its functions is the giving of vocational guidance to boys and girls who are about to leave school, and to young persons seeking employment. It also conducts psychological researches bearing on school life and vocational guidance[2]. Vocational and intelligence tests are also used in several denominational bureaux for vocational guidance, e.g., the Central Catholic Bureau at Utrecht.

Italy.

(A) *Tests of Intelligence.*

Tests of Intelligence are at present only used in some psychological laboratories and psychiatrical clinics, e.g. by Professor Ugo Pizzoli in the Instituto di Crevalcore near Modena and by Professor de Sanctis in the laboratory of experimental psychology attached to the Royal University of Rome.

The tests are also applied to defective children in several special schools and institutions for such children at Rome under the direction of Professor de Sanctis[3].

---

[1] From information supplied by Dr. Bokhorst of Amsterdam and Professor Brugmans of Groningen.

[2] From information supplied by Dr. van Wayenburg and Miss Kalshoven of Amsterdam.

[3] From information supplied by Professor de Sanctis.

(B) *Vocational Guidance.*

Vocational tests are at present being used experimentally by Prof. de Sanctis at Rome, by Prof. Ugo Pizzoli in the Instituto di Crevalcore, and by Professors Gemelli and Corberi at Milan[1].

The Societá Umanitaria of Milan has a psychotechnical laboratory in connection with its Labour Clinic, at which investigations in the use of vocational tests are carried out[2].

Researches in the use of vocational tests are also being conducted in the psychotechnical laboratory " Fr. Menafoglio " attached to the technical school " Fermo Corni " at Modena[3].

Japan.

The Department of Industrial Psychology, established at Tokio in 1921 conducts psychological research and investigates problems of vocational selection.

The Institute of Applied Psychology attached to the University of Tokio is constructing a series of tests for vocational aptitude[4].

Luxemburg.

The Aceries Réunies de Burdach-Eich-Dudelange "A.R.B.E.D." in Luxemburg, maintain a psycho-physiological laboratory in connection with the Institut Emile Metz, which is a Technical School for apprentices in the works. Psychological and physiological tests are applied to all apprentices both on entry and at periods during the course. Careful records are kept of the results obtained by the tests[5].

Norway.

A vocational organisation has been established at Christiania, the main work of which at present is the standardisation of tests that are to be used later for vocational guidance and for testing

---

[1] From information supplied by Professor de Sanctis.

[2] *Rivista della Psicologia*, Anno XIX (1923), p. 5.

[3] *Rivista della Psicologia*, Anno XIX (1923) pp. 97–99.

[4] *Journal of National Institute of Industrial Psychology*, Vol. I, p. 111 (for July, 1922). Fontègne, op. cit., pp. 140–141.
International Labour Review, Vol. V, p. 715 (for May, 1922).

[5] Cf. *A.R.B.E.D. Oeuvres Sociales*, Luxembourg, Imprimerie de la Cour, 1923, pp. 59–61.

officers for subordinate positions in the Navy and Army. The funds are provided by the State University and from private sources, and the naval and military work is paid for by the Ministry of Defence[1].

Spain.

(A) *Tests of Intelligence.*

Tests of intelligence are at present scarcely used at all in schools, though some attention is now being devoted to them in the Instituto-Escuela de Segunda Enseñanza at Madrid[2].

(B) *Vocational Guidance.*

The Institut D'Orientacio Professional, Calle San Honorato, Barcelona. This Institute which was founded in 1920 receives an annual grant of 20,000 pesetas from the City of Barcelona and a like grant of 20,000 pesetas from the Principality of Catalonia, equivalent to about £1,400 a year. The Director is assisted by a council of fourteen persons, including a chemist, an engineer, a schoolmaster and a schoolmistress, which meets weekly to review the work carried out in the Institute. The Institute which is only open in the evenings is divided into four departments :—

(1) The Medico-Anthropometric Department ;
(2) The Psychometric Department ;
(3) The Statistical Department ;
(4) The Information Department.

The expert heads of the two first-named departments are assisted by two young doctors. Children seeking vocational guidance are first subjected in the Medico-Anthropometric department to a thorough medical examination, including any details about the medical history of the family and so forth. In the Psychometric department tests are applied to measure the intelligence, judgment, reasoning powers, emotional reaction, reaction times, attention, memory, imagery and so forth of each applicant who is also required to reply on a form to printed questions regarding his various interests, ideals, desires, feelings, social environment, etc.

Teachers at the public elementary schools advise children of school-leaving age to visit the Institute. The teachers are provided with forms on which they transmit to the Institute the names of each

---

[1] From information supplied by Dr. Martin L. Reymont of Christiania.

[2] From information supplied by Señor José Castillejo of Madrid.

applicant with a brief summary of his school career and his mental and moral attainments. Attendance at the Institute is optional for school children, but their names cannot be placed on the books of certain labour bureaux until they have received advice from the Institute. Youths and girls attending the Technical College at Barcelona are required to undergo examination at the Institute if they are not making satisfactory progress in their studies. The Department of information records: (*a*) a resumé of the results obtained by the application of the laboratory tests to each applicant, his personal and family history and the advice tendered to him; (*b*) the economic and social conditions of about 50 occupations and the special physical and mental requirements of each; (*c*) the names and addresses of the various unions and federations of employers and employees together with the occupations included in each[1].

Sweden.

There is a Psychological Laboratory at Gothenburg which is engaged in collecting psychological tests and other material for the formation of a Vocational Bureau. Its income is derived from private sources with occasional grants from the Educational Authorities for printing and so forth[2].

Switzerland.

(A) *Tests of Intelligence.*

Tests of Intelligence are used in a few elementary and secondary schools. For some time past the Institut J. J. Rousseau at Geneva, the Pestalozzianum at Berne and the Psychological Institute at the University of Zürich have been conducting experiments in the application of tests of intelligence to school children.

(B) *Vocational Guidance.*

For several years past the Technical Schools at Lausanne, Chaux-de-Fonds and Geneva have been using tests of motor ability and mechanical aptitude as a part of their entrance examinations[3].

---

[1] From information supplied by Dr. C. S. Myers. See also *Anals de 'Institut d'Orientacio Professional*, Barcelona, 1920–1923, *passim.* and *Journal of National Institute of Industrial Psychology*, Vol. I, pp. 31–32 and p. 211.

[2] From information supplied by Professor G. A. Jaederholm, Gothenburg.

[3] From information supplied by Dr. E. Claparède and others. Cf. Fontègne, op. cit., p. 134.

Geneva, Institut J. J. Rousseau, Cabinet D'Orientation Professionelle. The work chiefly consists of inquiries in regard to occupations under the following heads :—

(1) The study of psychological tests and other means of guidance and selection.

(2) The selection of candidates for Technical and Trade Schools.

(3) The training of psychologists for work in vocational guidance and for psychotechnical consultation.

The Office, which was established in October, 1922, is maintained by grants from the Institut J. J. Rousseau, subscriptions, donations and consultation fees.

The practical work in the Laboratory consists of the construction and development of tests, the study of problems such as educational ability, movements, and the analysis of the physical and mental characteristics of various occupations. The giving of assistance in the practical work of vocational guidance leads up to direct collaboration with employment agencies[1].

Zürich, Psychotechnisches Institut. This Institute works independently, but receives support from the educational authorities of the Canton of Zürich, and is also closely connected with the University and with the chief local industrial concerns. Its income chiefly consists of fees for services rendered. The vocational work has a twofold aspect : (*a*) enquiries are made regarding the requirements of the principal occupations of the district. This work has already been carried out by the various local authorities, but the Institute has endeavoured to make a more thorough investigation from the psychological standpoint: (*b*) the Institute also examines young persons who desire to enter the various local occupations and gives a report to each candidate. From the information obtained by these examinations and from their knowledge of the state of the neighbouring industries and occupations, and their records of the requirements of each several occupation, the staff of the Institute are in a position to give advice as to what occupation a child should take up. The Institute also undertakes a systematic examination of the higher classes of the local Elementary Schools, with the double aim of selecting promising pupils for promotion or for further education, and for giving vocational guidance to those who will later be leaving School[2].

---

[1] From information supplied by Dr. G. H. Miles, cf. also Fontègne, op. cit., pp. 61–63, and pp. 132–134.
International Labour Review, Vol. V, p. 714. L'Education en Suisse, Geneva, 1914, pp. 104–105.

[2] From information supplied by Dr. G. H. Miles.

The United States of America.

(A) *Tests of Intelligence and Standardised Scholastic Tests.*

The use of the tests has not yet been made official by the Education Authorities of any one State, but those of many States have taken much interest in them and are developing their own tests and comparing the results with one another. For example, Wisconsin has a "supervisor of educational measurements," Maryland has been giving tests in the rural districts and providing demonstrations for the use of specialists; Michigan uses the tests in rural schools for the purpose of selecting children for special classes. The interest in the tests originated in the higher educational institutions such as Universities and Normal Schools and has spread downwards to the ordinary schools until in almost every State some form of the tests is now being employed. In one State, for example, rudimentary tests have been in use for ten years, and have resulted in appreciable improvement in the children's work. The National Research Council, consisting of the most distinguished psychological specialists in America, and established by voluntary effort as the outcome of the activities of the several States, has constructed and standardised a set of tests, known as the National Intelligence Tests. Another organisation due to private initiative is the Psychological Corporation, founded early in 1922, which includes among its Directors most of the leading psychologists in the United States[1]. One of its objects is to render expert services involving the application of psychology to educational problems and by research to improve and standardise tests to aid in the problem of selection for general intelligence and for special aptitudes.

In a memorandum sent to the Consultative Committee by members of the Department of Psychology at Teachers College, Columbia University, it is stated that "intelligence" tests[2] are used for the following purposes :—

(1) *In Special Classes for Backward or Mentally Dull Pupils.* In many school systems where such classes have been organised, selection for them is based entirely on a mental examination which is usually of the kind such as may be obtained by the use of the Binet Scale or some similar intensive method of examination. The tendency now in segregating children for such classes is to use group tests or other objective means of measurement rather than the judgment of the teacher, though the latter is never wholly neglected.

---

[1] *Journal of the National Institute of Industrial Psychology*, Vol. I, No. 2, p. 76.

[2] Cf *The Twenty-first Yearbook of the National Society for the Study of Education, Intelligence Tests and their use. passim.* Public School Publishing Co., Bloomington, Illinois.

(2) *In Special Classes for Exceptionally Bright Children.* In school systems where such classes have been organised up to the present, " intelligence " tests of some kind have generally been employed in the selection of children, and the use of such tests is stated to stimulate greatly the interest of teachers in the exceptionally gifted children.

(3) *In the General Classification of School Children.* " Intelligence " tests, especially group tests, are being more and more used to classify the ordinary children in schools so that pupils of like mental gifts may be brought together in one class. In many cases where there are several sections of the same class the division into sections is based on the results of " intelligence " tests, as it has been found that a short and scientifically not very accurate test often serves for a rough classification of pupils into two or more groups. Where this has been done, the teachers report favourably.

(4) *In Psychological Clinics.* Clinics for the more intensive study of individual cases have been established in connection with Universities, hospitals, penal and charitable institutions and also in some public school systems. These clinics make extensive use of all kinds of psychological and standardised educational tests in their study of individual cases, and those attached to the Universities train students in the use of such tests.

(5) *In Bureaux of Educational Research.* These Bureaux, developed in connection with various city school systems and with some State Departments of Education, make wide use of " intelligence " and standardised scholastic tests and supervise the application of such tests in the school system. Many of them carry on a certain amount of testing in schools as part of their regular routine work.

(6) *In Universities and Colleges.* Many institutions of academic rank use " intelligence " tests in addition to the usual entrance examinations. The results obtained by the application of intelligence tests to entrants at Columbia University are reported to have been particularly satisfactory[1]. Even Universities which do not apply tests as an entrance requirement nevertheless find the results of such tests valuable to assist them in the guidance of the students during their College career.

(7) *In the Army.* More than 1,000,000 soldiers were tested by psychologists during the Great War and the results obtained have stimulated research and led to a better understanding of group tests and their possibilities.

The tabular statement on the following pages, kindly prepared for the Committee by Dr. Mary Whitley, gives some idea of the extent to which intelligence tests are being used in municipal systems of Education throughout the country.

[1] See Prof. E. L. Thorndike's articles in *Journal of Educational Research* for May, 1920 and September, 1921.

City or town and school population at last census.	Proportion of schools using tests.	Is grading and promotion determined by tests ?	In what part of school are they applied ? When ?	Who administers the tests ?
Baltimore (Maryland), 109,840.	10% Elementary Schools. 80% Junior High Schools. 60% High Schools.	In part : achievement and school record are also taken into account.	In 20 schools in the 1st grade. In 12 schools throughout the school.	Specially interested teachers and principals. Committee of 3–6 persons specially trained.
Chicago (Illinois), 484,510.	50% The Illinois Intelligence Tests are used.	In part : the teacher's opinion and physical condition are also taken into account.	Several High Schools in entering classes.	Regular teachers or principal.
Cleveland (Ohio), 150,638	All schools	In part	Grades 1 to 6. 1 B in most schools. All entrants to High Schools.	Principals, or supervisory assistant. Specially qualified High School teachers. Teachers, examiners from H.Q. All under the Division of Research.
Denver (Colorado), 54,855.	All schools. The National Intelligence Tests are used.	Not in any fixed definite way.	At the end of the school year, for next year's classification. All high schools.	Special officers.
Kalamazoo (Michigan), 9,520.	All schools. The National Intelligence Tests are used.	In part: school record, teacher's estimate, life history, parents' wishes and social environment are also taken into account.	In entering grades at the beginning of the school year.	Special supervisors and experts, assisted by the regular teachers.

New Haven (Connecticut), 36,847.	20%	In part	To pupils on entering grades, and on promotion to higher grades.	Regular teachers, supervised by specially trained directors.
New Rochelle (New York), 8,650.	All schools	In part : individual tests are also used in certain cases.	In grades 3 to 6 mostly : some other classes. At beginning of year for grading and prognosis.	Specially trained teachers.
Oakland (California), 49,330.	All schools	In part	In the 4th, 6th and 8th grades.	Bureau of research, or certificated principals and teachers.
Oswego (New York), 2,927.	In several schools ..	Used as a check.. ..	High School entering classes.	Principal.
Passaic (New Jersey), 14,390.	Used once, in the largest Elementary School.	—	Especially in the kindergarten and first grade.	Supervisors and principal.
Philadelphia (Pennsylvania), 314,308.	All schools promoting to Junior or Senior High School.	In part	Entering grades at beginning of school year, for classification and prognosis.	School principals and teachers under their direction.
Rochester (New York), 218,150.	70%	Yes, teacher's estimate is also taken into account.	6th and 8th grades. Before they go to Junior High or High School : all entering classes in High School.	Special officers, never the regular teachers.

(B) *Vocational Guidance.*

The Bureau of Vocational Guidance at Harvard University, founded in 1917, is probably typical of a large class of academic institutes for vocational guidance and vocational selection. The Bureau gives information to many teachers and others who write for information about vocational guidance, and is frequently consulted by individuals requiring vocational advice. Similar instruction in Research on Vocational Psychology is carried on in most of the Universities and other institutions of academic rank in the United States, *e.g.*, Columbia University, New York, Bryn Mawr, Chicago, Michigan, Yale, the Carnegie Institute of Technology at Pittsburgh[1].

The American Vocational Guidance Association has adopted a number of principles for vocational guidance, from which we quote the following passage :—

> " For all children before the school-leaving age there should be provided a wide variety of try-out experiences in academic and æsthetic work, gardening, simple processes with tools and machines, elementary commercial experiences, and co-operative pupil activities. Such try-out experiences are for the purpose of teaching efficiency in every-day tasks, broadening the social and occupational outlook of the children, and discovering to them and the teachers their interests and abilities.
>
> " Tests of general intelligence should be used with the greatest care. No important decision should be made on the basis of a group test alone ; special classifications and assignment of special curriculums should be made only after an individual examination by a carefully trained and experienced psychologist. Whenever time and facilities permit, occupational tests should also be used.
>
> " Cumulative records should be kept for individuals. These should include academic records, social conditions, physical and mental records, and the results of counselling.
>
> " School systems should undertake follow-up work and employment supervision, to extend throughout the time of minority of the child and to be exercised in co-operation with the above-mentioned agencies."

Organisations for vocational guidance exist in most of the large cities in the United States and exhibit all stages of development.

The Director-General of the United States Department of Labour recently wrote :—

> " There is no uniform and complete system of vocational guidance under federal direction in the United States. The movement known as vocational guidance ' has developed

[1] Journal of the *National Institute of Industrial Psychology*, Vol. I, No. 1, p. 28, and Vol. I, No. 4, p. 152. Fontègne, op. cit., pp. 114–123.

locally and sporadically throughout the country. However, through the Junior Division of the United States Employment Service certain centralised and system work is being done which encourages its development and tends to standardise methods.

"The Junior Division in question conducts experiments in the field of vocational guidance and in the placing of children and young persons, and aims at formulating general policy and collecting information in order to serve in a national clearing house for information. Grants are made for work undertaken locally."

A good example of municipal organisations for vocational guidance is the Bureau of Research and Guidance of the public schools at Oakland, California, which provides a vocational adviser in every large elementary school and in every high school of the district. It includes the departments of research, vocational guidance, placement, industrial welfare and mental testing.

So, too, the Pittsburgh public schools have an extensive Vocational Guidance and Placement Department which co-operates with the municipal Junior Placement Office. Another typical organisation of the sort is "Vocational Guidance for Juniors Incorporated," New York.

The Children's Bureau established at Washington in 1912 makes a special study of methods including tests for the guidance and placing of boys and girls and has undertaken a survey of existing methods in some 15 typical cities.

APPENDIX IV.

Note by Dr. Cyril Burt on Standardisation and Norms.

The work of "standardisation" as applied to psychological tests has two sub-divisions: (1) standardisation of method, (2) standardisation of results.

(1) *Standardisation of method.* It is a fundamental principle of scientific testing that for purposes of exact comparison the method employed must be the same for all examinees, for all examiners, and for all different occasions. This requires not only that the material (sums, problems, sentences, apparatus, pictures) should be essentially the same in all cases, but that the questions should always be phrased in the same terms. Teachers unfamiliar with the technique of psychological testing sometimes modify their questions freely when they repeat them, in the hope of rendering them more intelligible to the child: an essential feature of psychological method is that the formulæ are practically invariable. Indeed, Binet expressly states that the novelty of his scale consists not in the nature of the questions, but in the fact that they are always set in the same way[1].

[1] Binet & Simon—Article on "Le Developpement de l'Intelligence," in *L'Année Psychologique* (1908), p. 60.

Here, however, one reservation must be made. The similarity of conditions required must be not objective, but subjective. In dealing with children of different temperaments the examiner's attitude must be correspondingly modified or adapted so that the conditions may really, so far as possible, be the same for all. If the same test is repeated time after time with the same children, it is clear that the content should be varied so as to minimise the effect of use and familiarity. For this purpose it is therefore desirable to obtain alternative forms of equal difficulty, *e.g.*, in the " Opposites " test, where it is perfectly practicable to draw up half a dozen different sets, each containing fifty problems and each being of the same difficulty as the others[1].

(2) *Standardisation of results.* For most purposes it is necessary to compare the examinees not only among themselves, but in relation to a fixed standard. If this fixed standard represents an ideal towards which the examinee's work is expected to approximate, but which it is not expected to attain, it is a standard *for* achievement. If it represents the actual level of a normal class or group, from which there can be divergences above and below, it is a standard *of* achievement. The former of these are of more interest to teachers, who prefer to watch the progress of their pupils towards an external ideal; the latter are of more interest to psychologists, who are more concerned with facts than with aims. At a meeting of the (American) National Association of Directors of Educational Research it was recommended that standards *for* achievement should retain the name of standards, and that those *of* achievement should be distinctively called norms.

Norms. By norms, in this sense, are meant specimens of work which represent the commonest type of achievement for the whole group in question. They constitute the means by which can be measured the degrees of abnormality shown by examinees above and below the normal. Hence they require at least two criteria: (1) a measure of the central tendency or average, (2) a measure of deviation.

The central or normal tendency for any given group is best measured by the average or arithmetical mean, which usually coincides with the mid-point between the extremes of abnormality—

[1] A test is commonly said to be " reliable " if it agrees consistently with itself, *i.e.*, if, when applied on successive occasions it yields approximately the same result with the same group. It is said to be " efficient " if it agrees with some external criterion accepted as valid, *e.g.*, the estimate of a competent and conscientious observer. The measure of agreement (usually expressed by a coefficient of correlation) shows the success with which the test has been standardised.

a 50 mark between zero and 100. It is most easily estimated when the work can be quantitatively marked. When the measurements are not in the form of quantitative marks, but are based on specimens or samples, *e.g.*, in drawing or handwriting, the best way of identifying it is by using a "median" or middle specimen.

To measure deviation a standard is required by reference to which the examiner can say not only that a given child is above or below the average, but by how much he exceeds or falls short of it. Binet, for example, measures the average in terms of mental ages, and the deviation from that average in terms of backwardness or advancement of one or more mental years. This is perhaps the simplest of all measures of deviation; others which are also in use will be found described in the numerous handbooks dealing with the application of statistical methods to education and pyschology[1].

APPENDIX V.

Note by Dr. Cyril Burt on Correlation as Applied to Mental Testing.

The statistical device known as correlation is widely used by education psychologists for measuring the validity of their methods. To test a given test, it is necessary to compare the results obtained from large and representative groups of children, first with later applications of the same test, and secondly with an independent criterion, such as the estimate of a competent teacher who knows intimately all the members of the group. A coefficient of correlation is an index-number intended to measure, on a scale of 0 to 1, the amount of agreement between two such series of estimates. Where the agreement is perfect, the figure is unity and positive, (+1.00); where there is no agreement whatever, the correlation is zero (0·00); where disagreement is at a maximum, one estimate exactly reversing the order of the other, the coefficient is negative (−1·00). Where agreement is more or less imperfect, the coefficient is a fraction, ranging between these extremes.

The following hypothetical instances illustrate both the nature of correlation generally and what is perhaps the most precise and practicable formula for calculating it in a given case. (See Table I).

---

[1] Useful references both on the standardisation of tests and upon the definition and use of norms are Whipple's Handbook, 2nd ed., first three chapters, and McCall's *How to Measure in Education*, Part II., "How to Construct and Standardise Tests" (see Chapter 11 for "Determination of Reliability and Norms.")

Suppose, for simplicity, there is a small class of ten children (named A.B., C.D., etc., see col. 1), ranked in order of intelligence by the teacher's estimate from 1st to 10th (col. 2). (In what follows this order will be assumed to be the final criterion.) Let us further suppose that to these children some single test is applied—for example a scholarship examination—which gives for the same class a second ranking in order of intelligence. The problem is: how closely does the second order correspond with the first, and how can we measure different degrees of such correspondence? Clearly, the lack of correspondence can readily be gauged by counting up the total number of discrepancies between the two rankings. The fewer the discrepancies, the higher will be the measure of agreement.

The absolute number of differences in ranking must tend, of course, to increase with the size of the group. We can allow for this by first determining for a group of this size the maximum number of differences in ranking obtainable, or (more simply) the probable number of discrepancies to be expected by pure chance. The proportion between the *actual* number of discrepancies and the *expected* number of discrepancies measures the amount of disagreement; and, by subtracting this proportion from unity, we obtain a measure of positive agreement. (See formula at foot of table.) Professor Spearman's simple "foot rule" formula does, in fact, consist in this simple calculation.

(1) Suppose, first, (example A) that the ranking supplied by the test (col. 3) is identical with that supplied by the teacher (col. 2). There will be no discrepancies (col. 4). The sum total of differences in ranking will be zero. It is clear that the formula proposed will at once yield a coefficient of $+1\cdot00$. The correlation is positive and complete.

(2) Suppose now (example B) that the test-result bears no relation whatever to the result of the teacher's estimate. It is a purely random order (col. 5). Here, for instance, the order given was obtained by shaking up ten dominoes in a dice-box, and drawing them one by one, as it were, by lot. Clearly, the discrepancies so obtained should be almost identical with the number of discrepancies that we might expect between any two chance arrangements. It will be seen that the nature of the formula yields a coefficient of approximately zero.

(3) Thirdly, suppose (example C) that the test exactly reverses the order of the teacher, so that the child whom the teacher put at the top of the list is placed by the test at the bottom, and the child whom the teacher placed last is placed first by the test. It is clear that the discrepancies will here be a maximum. And the application of the formula necessarily yields a coefficient of $-1\cdot00$.

[In the formula proposed the only difficult portion is to calculate what, for a group of a given size, is the total number of differences in ranking to be expected by pure chance. This is deduced from the mathematics of probability. (It is best, first of all, to square all the rank-differences—a well-known algebraical device for eliminating the differences of plus and minus signs (cols. 7, 10, 13). This squaring constitutes the main difference between the present formula and the simpler "foot-rule" of Professor Spearman.) The non-mathematical can readily understand the calculation by looking more closely at the extreme case of negative correlation (example C). He will perceive that the maximum number of differences obtainable is, by the very nature of the computation to be made, exactly equal to *twice* the sum of the alternate numbers from 0 to the number just before the total number in the group, *i.e.*, since the group contains ten children, the sum of the odd numbers from 1 to 9 (col. 9).

It will be readily understood that the actual number of differences in ranking (squared) to be expected by chance is exactly midway between the number of such differences obtained with complete positive correlation, namely, zero, and the number of such differences expected with complete negative correlation; that is to say, it is exactly *once* the sum of squares of the aforesaid alternate numbers. The algebraic fraction which appears in the denominator of the correlation-formula is, in fact, the well-known formula for calculating the sum of such squares].

(4) Coefficients of exactly unity or exactly zero are rare. The kind of ranking actually obtained in such an examination is illustrated by the ranking given in example D (col. 11). Here it will be seen that the children marked $1\frac{1}{2}$ and the children marked 5 are "ties"; and have therefore been given the midway ranking. Applying the formula just deduced, the coefficient proves to be about ·7.

With an actual test this would be considered rather a high degree of correlation. As a matter of fact, however, the ranking given in this illustration (col. 11) was obtained by averaging the teacher's estimate (col. 2) with the chance order (col. 5), thus giving both an equal weight. The distribution of the share of each is, as it were, fifty-fifty. The uninstructed teacher is apt to assume that, with a correlation of ·7, the part played by chance is much smaller, in the ratio of (1–·7) to ·7, that is 3 parts out of 10. As a matter of fact, the coefficient must first be squared; and the proportion of chance and of intelligence is as $(1{\cdot}00 - {\cdot}71^2)$ to $\cdot 71^2$, that is (since $\cdot 71^2 = \cdot 5$ approx.) 5 parts out of 10.

Thus, if the correlation coefficient is the best measure of agreement between two series, the square of the coefficient is the best measure of the part contributed by the factor common to both.

TABLE I.—THE NATURE AND CALCULATION OF COEFFICIENTS OF CORRELATION (RANK-METHOD).

		(A) *Perfect Correlation.*		(B) *No Correlation.*			(C) *Negative Correlation.*			(D) *Positive but Imperfect Correlation.*		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Name.	Teacher's Estimate.	Ranking for Test.	Rank-differences.	Ranking for Test.	Rank-differences.	Rank-differences squared.	Ranking for Test.	Rank-differences.	Rank-differences squared.	Ranking for Test.	Rank-differences.	Rank-differences squared.
		(*a*)		(*b*)			(*c*)			(*d*)		
A.B.	1	1	0	6	+ 5	25	10	+ 9	81	3	+ 2	4
C.D.	2	2	0	3	+ 1	1	9	+ 7	49	1½	− ½	¼
E.F.	3	3	0	2	− 1	1	8	+ 5	25	1½	− 1½	2¼
G.H.	4	4	0	7	+ 3	9	7	+ 3	9	5	+ 1	1
I.J.	5	5	0	10	+ 5	25	6	+ 1	1	8	+ 3	9
K.L.	6	6	0	5	− 1	1	5	− 1	1	5	− 1	1
M.N.	7	7	0	9	+ 2	4	4	− 3	9	9	+ 2	4
O.P.	8	8	0	4	− 4	16	3	− 5	25	7	− 1	1
Q.R.	9	9	0	8	− 1	1	2	− 7	49	10	+ 1	1
S.T.	10	10	0	1	− 9	81	1	− 9	81	5	− 5	25
Total of differences..			0		32	164		50	330		18	48½

Formula: $r = 1 - \frac{\Sigma(d^2)}{\frac{n(n^2-1)}{6}}$	$r = 1 - \frac{0}{165}$ $= 1\ 0$ $= 1 \cdot 00$	$r = 1 - \frac{164}{165}$ $= 1 - 1$ (approx.) $= 0 \cdot 00$ (approx.)	$r = 1 - \frac{330}{165}$ $= 1 - 2$ $= -1 \cdot 00$	$r = 1 - \frac{48 \cdot 5}{165}$ $= 1 - \cdot 29$ $= \cdot 71$

Coefficient of Correlation $= 1 - \frac{\text{Sum of actual rank differences (squared)}}{\text{Sum of rank differences (squared) expected by chance.}}$

APPENDIX VI.

Note, by Mr. A. E. Twentyman, on Grades in American Schools.

In America, as in England and Wales, the earliest form of elementary school organisation was the group of children under a single teacher. When the growth of population and the increase of material resources made possible a more extended system, the plan generally adopted was to establish two practically independent schools each with its own teacher and separate course of study. Frequently such schools were conducted on different floors of the same building, and were known as the Primary School and the English Grammar School. At a later stage another school—variously called the Secondary or Intermediate School—was interposed between the Primary and the Grammar School. The term "graded school" was in common use in the 'fifties, but it referred to this grouping of schools rather than to the classes within the school.

About 1865 the cities began to organise their schools under the direction of a city superintendent, and schools of the various grades were amalgamated to form a single school under a principal. About this time the term "grade" came to signify a stage in the school course, but the precise meaning attached to it differed with the locality. While the "grade" usually stood for a year's work, in New York City there were ten half-yearly grades included in the Primary Course. No supreme education authority existed which could bring order out of this confusion. In 1874, at the meeting of the National Education Association, the school superintendents adopted a conventional grading system. The elementary school was divided into two departments, the Primary School Department and the Grammar School Department, each containing four progressive years or grades of instruction.

This form came to be adopted as the normal type of school organisation, but the conditions varied so considerably that Grade VIII by no means always represented the same standard of achievement. In some States school attendance began at 6 years of age, in others at 7, and in many at 8. In addition, the compulsory period of attendance in a year varied within very wide limits. Thus the "grade" had a far less stable meaning than the old "standard."

Moreover, in the last twenty years there has been a constant stream of protests against the rigidity of the grade system with its annual promotions. In some cities the courses are so organised that the brighter children complete in six years what the average child achieves in eight. In others, there are three parallel courses throughout the grades in which the children progress at different rates. The content of the course is fuller for the abler child than for the average, and of the dull child less than the average course is expected.

A more complete departure from the grade system is represented by the attempts to "departmentalise" the last two grades—i.e., to introduce what we should call the set system, allowing the pupils to be reclassified according to their proficiency in each subject. The opportunity is also afforded to introduce other subjects than those usually taught in the present grades. There is a growing demand for a reorganisation of the whole scheme. It is suggested that if waste be eliminated, the present programme of the elementary school could be achieved in six years, and that this shortened course should be followed by six years of secondary education. In many localities this has been achieved by the establishment of Junior High Schools. In such places the old grades 7 and 8 have practically disappeared.

APPENDIX VII.

A Short List of Recent Publications on Psychology and Psychological Tests[1].

(Prepared for the Committee by Dr. Cyril Burt.)

(A) General Psychology.

McDougall, William. *Psychology : The Study of Behaviour.* Home University Library. Williams and Norgate. 1912. 2*s.* 6*d.*

McDougall, William. *An Outline of Psychology.* Methuen & Co. 1922. 12*s.*

Woodworth, R. S. *Textbook of Psychology.* Methuen & Co. 1922. 8*s.* 6*d.*

Stout, G. F. *Manual of Psychology.* University Tutorial Press. 1904. 10*s.* 6*d.*

James, W. *Principles of Psychology.* Two vols. Macmillan & Co 1901. 31*s.* 6*d.*

Myers, C. S. *Textbook of Experimental Psychology.* Two vols. Cambridge University Press. 1911. 12*s.*

Spearman, C. *The Nature of Intelligence and Principles of Cognition.* Macmillan & Co. 1923. 15*s.*

Brierley, S. S. *An Introduction to Psychology.* Methuen. 1921. 5*s.*

British Journal of Psychology (General Section). Cambridge University Press. Quarterly. 5*s.* to 9*s.*, according to size of number.

---

[1] Dr. Burt desires to acknowledge the help he has received from psychological and educational colleagues in the preparation of this bibliography for the Committee.

(B) Educational Psychology.

Nunn, T. P. *Education: Its Data and First Principles.* Arnold. 1920. 6*s.*

Rusk, R. *Experimental Education.* Longmans, Green & Co. 1919. 7*s.* 6*d.*

Thorndike, E. L. *Educational Psychology.* Three vols. Kegan Paul. 1913–14. 63*s.*

Drever, J. *An Introduction to the Psychology of Education.* Arnold. 1922. 6*s.*

Starch, D. *Education Psychology.* The Macmillan Co. 1910 14*s.*

Fraser, D. Kennedy, *Psychology of Education.* Methuen. 1923. 6*s.* 6*d.*

The Forum of Education. Longman, Green & Co. Three times a year. 1*s.* 6*d.* per number; 5*s.* per annum.

Journal of Educational Psychology. Warwick & York. Monthly, except July and August. 50 cents. per number; $3.00 per annum.

Journal of Educational Research. Public School Publishing Co., Bloomington (Ill.). Monthly, except July and August. $4.00 per annum.

The Pedagogical Seminary. F. Chandler, Worcester (Mass.). Quarterly. $1.50 per number; $5.00 per annum.

(C) Mental Measurement and Statistics.

McCall. *How to Measure in Education.* Macmillan & Co. 1922. 15*s.*

Rugg, H. O. *Statistical Methods Applied to Education.* Harrap & Co. 1922. 10*s.* 6*d.*

Brown, W., and Thomson, G. *The Essentials of Mental Measurement.* Cambridge University Press. 1921. 20*s.*

Yule, G. U. *Introduction to the Theory of Statistics.* C. Griffin & Co. (5th ed.) 12*s.* 6*d.*

(D) General Mental Testing.

Ballard, P. B. *Mental Tests.* Hodder & Stoughton. 1920. 6*s.*

Whipple, G. M. *Manual of Mental and Physical Tests.* Warwick and York. (2nd ed.) 1914–15. 25*s.*

Stern, W. *The Psychological Methods of Measuring Intelligence.* Translated by G. M. Whipple. Warwick and York. 1913.

Gregory, W. *Introduction to Mental Measurement.* Appleton. 1923. 8*s.* 6*d.*

National Society for the Study of Education. *Twenty-first Yearbook: Intelligence Tests and their Use.* Public School Publishing Co, Bloomington (Ill.). 1922.

(E) The Binet-Simon Scale.—Instructions for Giving the Binet-Simon Tests.

Binet, A., and Simon, Th. *The Development of Intelligence in Children.* Translated by Elizabeth S. Kite. Publications of the Vineland Training School, New Jersey. 1916. $2.00.

Binet, A., and Simon, Th. *A Method of Measuring the Development of Intelligence in Young Children.* Translated by Clara H. Town. Chicago Medical Book Co. 1913. Revisions of the Binet-Simon Tests.

Burt, C. *Mental and Scholastic Tests.* (L.C.C. Reports, 2052.) P. S. King & Son. 1921. 21*s.*

Terman, L. M.; Lyman, G.; Ordahl, L.; Galbreath, N.; and Talbert, N. *The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence.* Warwick & York. 1917. 7*s.* 6*d.*

Terman, Lewis M. *The Measurement of Intelligence.* G. G. Harrap & Co. 1919. 8*s.* 6*d.* (Test Material, 3*s.* 6*d.*)

Terman, Lewis M. *Intelligence Tests and School Reorganisation.* Harrap & Co. 1923. 4*s.* 6*d.*

Yerkes, R. M.; Bridges, J. W. Hardwick; Rose, S. *A Point Scale for Measuring Mental Ability.* Warwick & York. 1915.

Melville, N. J. *Testing Juvenile Mentality.* Lippincott, C. 1917. 12*s.* 6*d.*

Herring, J. P. *Herring Revision of the Binet-Simon Tests: Examination Manual.* Harrap & Co. 1923. 5*s.*

(F) Group Tests of Intelligence.

Ballard, P. B. *Group Tests of Intelligence.* Hodder & Stoughton. 1922. 6*s.*

Northumberland Mental Tests (Godfrey Thompson). Specimen sets. Harrap & Co. 1*s.* 6*d.*

Simplex Group Intelligence Scale (C. A. Richardson). Harrap & Co. 3*s.*

National Institute Group Tests (C. Burt). Series 31, 32, or 33. National Institute of Psychology. 2½*d.*

Otis' Group Intelligence Scale. Primary Examination (Forms A and B), for grades 1–4. Advanced Examination (Forms A and B), for grades 5 through high school. With Directions and Key. Harrap & Co. 3*s.* 6*d.*

Terman, W. *Group Test of Mental Ability* (Forms A and B). For Secondary Schools. For grades 7–12. Specimen set 15 cents postpaid. Harrap & Co. 2*s.*

National Intelligence Tests. Scale A and B. (Forms 2.) With Manual of Directions and Key. (Haggerty, Terman, Thorndike, Whipple, and Yerkes.) For grades 3–8. Harrap & Co. 2*s.* 6*d.*

Yoakum, C. C., and Yerkes, R. N. *Mental Tests in the American Army*. Sidgwick & Jackson. 1920. 6*s*.

Whipple's Group Test for Grammar Grades. For grades 4–8. G. M. Whipple, University of Michigan, Ann Arbor, Mich.

(G) Standardised Scholastic Tests.

(*a*) *General.*

Starch, D. *Educational Measurements*. The Macmillan Co. 1916. 8*s*.

Monro, W. S.; De Voss, J. C.; and Kelly, F. S. *Educational Tests and Measurements*. Constable & Co. 1917. 6*s*.

Stanford Achievement Tests. (Terman, Kelly & Ruch.) World Book Co. 1922.

Burt, C. *The Distribution and Relations of Educational Abilities*. (L.C.C. Reports, No. 1868, 1919.) P. S. King & Sons. 2*s*. 6*d*.

Monroe, W. S. *Measuring the Results of Teaching*. Harrap & Co.

Pressey, S. L. & L. C. *Introduction to the Use of Standard Tests*. World Book Co. 1922.

Wilson, G. M. & Hoke, K. J. *How to Measure*. Macmillan & Co. 1920. 12*s*.

Burt, C. *Handbook of Tests for Use in Schools*. 1923. P. S. King & Son. 3*s*. 6*d*.

Ballard, P. B. *The New Examiner*. Hodder & Stoughton. 1923. 6*s*.

(*b*) *Reading.*

Thorndike's Scale Alpha 2 for Measuring the Understanding of Sentences. Part I, for grades 3–5; Part II, for grades 6–12. Teachers College, Columbia University, New York City.

Thorndike-McCall Reading Scale for the Understanding of Sentences. For grades 2–8. Teachers College, Columbia University, New York City.

(*c*) *Spelling.*

Thorndike's Visual Vocabulary Scale. For grades 3–10. Teachers College, Columbia University, New York City.

Haggerty's Achievement Examination in Reading. Sigma 1, for grades 1–3; Sigma 2, for grades 3–8; Sigma 3, for grades 5–12. Harrap & Co.

Ayres' Spelling Scale. Russell Sage Foundation, New York City.

Buckingham's Extension of the Ayres' Spelling Scale. For Elementary Schools. Public School Publishing Company, Bloomington (Ill.) 8 cents.

(*d*) *Arithmetic and Mathematics.*

Rogers, Agnes L. *Tests of Mathematical Ability.* Teachers College, Columbia. 1921.

Courtis' Standard Practice Tests in Arithmetic. For grades 4–8. (Includes the Courtis' Standard Research Tests and the Supervisory Tests in Arithmetic.) Harrap & Co. 5*s.*

Starch's Arithmetic Tests. For grades 3–8. The University Co-operative Co., 504, State Street, Madison, Wis.

Monroe's Diagnostic Tests in Arithmetic. For Elementary Schools. Kansas State Normal School, Emporia, Kansas.

Monroe's Standard Reasoning Test in Arithmetic. For Elementary Schools. Kansas State Normal School, Emporia, Kansas.

Woody's Arithmetic Scales. Two series for grades 2–8. Teachers College, Columbia University, New York City.

Thurstone's Arithmetic Examination for High School Graduates, with Problems appealing to Engineering Interest. For Secondary Schools and Colleges. L. L. Thurstone, Carnegie Institute of Technology, Pittsburgh, Pa.

Thorndike's Algebra Test. Reference: Mathematics Teacher. March, 1924.

Thurstone's Algebra Examination. For Secondary Schools and Colleges, especially engineering students. L. L. Thurstone, Carnegie Institute of Technology, Pittsburgh, Pa.

Thurstone's Constructive Geometry Test. For Secondary Schools and Colleges, especially for engineering students. L. L. Thurstone, Carnegie Institute of Technology, Pittsburgh, Pa.

(*e*) *English Composition and Grammar.*

Thorndike's English Compositions : 150 specimens arranged for use in Psychological and Educational Experiments. Teachers College, Columbia University, New York City. 1916.

Hillegas' Scale for Measurement of English Composition by Young People. For Elementary Schools. Teachers College, Columbia University, New York City.

Harvard-Newton Composition Scale (*Ballou*). For Elementary and Secondary Schools. Harvard University Press, Cambridge 38, Mass.

Thorndike's Extension of the Hillegas Scale. For Elementary and Secondary Schools. Teachers College, Columbia University, New York City.

Buckingham's Grammar Test. Third Conference on Educational Measurements, University of Indiana, Bloomington, Ill.; University of Illinois, Urbana, Ill.

Starch's English Grammar Tests. Three tests for grades 7–12. University Co-operative Company, 504, State Street, Madison, Wis.

Boyd, W. *Measuring Devices in Composition, Spelling and Arithmetic.* Harrap. 1924. 4*s.* 6*d.*

(*f*) *Handwriting.*

Ayres' Handwriting Scale for School Children. For Elementary Schools. Russell Sage Foundation, New York City.

Ayres' Measuring Scale for Handwriting. Three slant edition, for Elementary Schools. Russell Sage Foundation, New York City.

Thorndike's Handwriting Scale. For grades 5–8. Teachers College, Columbia University, New York City.

Courtis' Standard Practice Tests in Handwriting. (Courtis-Shaw (includes Courtis') Standard Research Tests in Writing and Courtis' Standard Supervisory Tests in Writing.) Harrap & Co. 45 cents.

(*g*) *Drawing.*

Rugg's Scale for Measuring Freehand Lettering. For use in Secondary Schools and Colleges. Journal of Educational Psychology, Vol. VI. January, 1915.

Thorndike's Drawing Scale for Grades 5–8. Teachers College, Columbia University, New York City.

Whitford's Art Tests. Test I.—Appreciation of the quality of Art; Test 2.—Ability to Draw. The Prang Co., Chicago, Ill.

(*h*) *Science.*

Thurstone's General Technical Information Test. L. L. Thurstone, Carnegie Institute of Technology, Pittsburgh, Pa.

Thurstone's Physics Problems. L. L. Thurstone, Carnegie Institute of Technology, Pittsburgh, Pa.

Ruch's Range of Information Test in General Science. G. M. Ruch, University of Oregon, Eugene, Ore.

(H) Practical, Manual and Performance Tests.

Healy, W., and Fernald, G. M. *Tests for Practical Mental Classification.* Psychological Monographs, XIII. 1911.

Pintner, R., and Paterson, D. G. *Performance Tests.* Warwick & York. 8*s.* 6*d.*

Dearborn, W. F.; Shaw, E. A.; and Lincoln, E. A. *Form-Board and Performance Tests of Intelligence.* Harvard Monographs in Education. Series I. No. iv. 1923.

Porteus, S. D. *Tests for Mental Deficiency: Vineland Revision.* C. H. Stoelting Co., 3037, Carroll Av., Chicago, Ill.

Kelly, A. *Constructive Ability Tests.* C. H. Stoelting Co., 3037, Carroll Av., Chicago, Ill.

Kohs, S. C. *Intelligence-Measurement based upon Block-design Tests.* Macmillan & Co. 1923. 14*s.*

Stenquist, F. *Tests for Mechanical Ability.* C. H. Stoelting Co., 3037, Carroll Av., Chicago, Ill.

(J) Vocational Psychology.

Claparède, E. *Problems and Methods of Vocational Guidance.* National Labour Office, Geneva. 1922. 2*s.*

Hollingworth, H. *Vocational Psychology.* Appleton & Co. 1920. 15*s.*

Munsterberg. *Psychology and Industrial Efficiency.* Constable & Co. 1918. 7*s.* 6*d.*

Myers, C. S. *Mind and Work.* Putnam & Sons. 1921. 3*s.* 6*d.*

Muscio, B., and Others. *Lectures on Industrial Administration.* Pitmans. 1920. 6*s.*

Muscio, B. *Review of the Literature on Vocational Guidance.* Reports of the Industrial Fatigue Research Board, No. 12. H.M. Stationery Office. 1921. 1*s.*

Link, H. C. *Employment Psychology.* Macmillan. 1919. 12*s.*

Thurstone, L. L. *Public Personnel Studies.* Bureau of Public Administration, Washington. 1922–24. 25 cents per number.

Fontègne, Julien. *L'Orientation professionelle.* Delachaux et Niestle, Paris. 1923.

Journal of the National Institute of Industrial Psychology. National Institute of Industrial Psychology. 21*s.* per annum.

(K) The Analysis and Testing of Temperamental Qualities.

Shand, A. F. *The Foundations of Character.* Macmillan Co. 20*s.*

Webb, E. *Character and Intelligence.* British Journal of Psychology. Monograph I. Cambridge University Press. 1912. 5*s.*

Jung, C. C. *Studies in Word Association.* William Heinemann. 1918. 25*s.*

The Personnel System of the U.S. Army. 2 vols. Govt. Printing Office, Washington.

Whateley-Smith, W. *The Measurement of the Emotions.* Kegan Paul. 1922. 10*s.* 6*d.*

Pressey, S. L. *X-O Tests.* Stoelting Co. 1922.

Downey, J. *Will-Temperament Test.* University of Wyoming, Laramie (Wyo.).

Green, G. H. *Psycho-Analysis in the Classroom.* University of London Press. 1923. 6*s.*

Flügel, J. C. *The Psycho-Analytic Study of the Family.* International Psycho-Analytic Press. 1921. 12*s.*

Slawson, J. *Reliability of Judgment of Personal Traits.* M. A. Thesis, Columbia. 1920.

Allport, F. H. & G. W. *Personality Traits: Their Classification and Measurement.* J. Abn. & Soc. Psychology. April, 1921.

Yerkes & La Rue. *Outline for the Study of the Self.* Harvard University Press.

(L) The Technique of the Oral Examination and Personal Interview.

Hollingworth. *Judging Human Character* (esp. Chap. V, "The Personal Interview"). Appleton & Co. 1922.

Rugg, H. O. "Is the Rating of Human Character Practicable." *Journ. Educ. Psychol.* Nov., 1921, Feb., 1922.

Knight and Franzen. "Pitfalls in Rating Schemes." *Journ. Educ. Psychol.* 1922.

Scott, W. D. "Scientific Selection of Salesmen." *Advertising and Selling Magazine.* Oct., 1915.

Board of Education. *Report of the Consultative Committee on Examinations in Secondary Schools.* London. 1911. Esp. pp. 74, 76, 155, 157, 365.

Hartog, P. J. *Examinations and their Relation to Culture and Efficiency.* London: Constable & Co. 1918. Esp. pp. 20–22; also pp. 52, 75, 85–86, 91.

Paulsen, F. *Prüfungen.* Article in Rein's "Encyklopädisches Handbuch der Pädagogik." 2te. Auflage. Vol. 7. Langensalza: H. Bryer & Söhne. 1908. (§ 3. Formen der Prüfung: Discusses the merits of written and oral examinations.)

Leathes, Sir Stanley. *What is Education?* London: G. Bell & Sons. 1913. Ch. VIII, pp. 159–161.

Board of Education. *Special Reports on Educational Subjects.* Vol. 24, pp. 287–289. (Account of the oral examination for the baccalauréat in France).

(M) Backward and Defective Children.

Tredgold, A. F. *Mental Deficiency.* Baillière, Tindall, & Cox. 3rd ed. 1920. 25*s.*

Binet, A., and Simon, Th. *The Intelligence of the Feebleminded* Translated by Elizabeth S. Kite. Publications of the Vineland Training School, New Jersey. 1916.

Binet, A., and Simon, Th. *Mentally Defective Children.* Translated by W. B. Drummond. Edward Arnold. 1914. 2*s.* 6*d.*

Goddard, H. H. *Feeblemindedness: Its Causes and Consequences.* The Macmillan Co. 1913. 27*s.*

Huey, E. B. *Backward and Feebleminded Children.* Warwick & York. 1912. 7*s.* 6*d.*

Woodrow, H. *Brightness and Dullness in Children.* Lippincott. 6*s.*

Bronner, Augusta F. *The Psychology of Special Abilities and Disabilities.* Kegan, Paul, Trench, Trubner & Co. 1919. 10*s.* 6*d.*

Porteus, S. D. *Studies in Mental Deviations.* Publications of the Vineland Training School. 1922. 13*s.*
Burt, C. *Backward Children* (L.C.C. Development Memoranda, No. 1). P. S. King & Son. 2*d.*
Studies in Mental Inefficiency. Central Assoc. for Mental Welfare. Quarterly. 10½*d.* per number; 3*s.* 6*d.* per annum.

(N) Juvenile Delinquency.

Healy, William. *The Individual Delinquent.* W. Heinemann. 1915. 21*s.*
Miner, J. Burt. *Deficiency and Delinquency: an Interpretation of Mental Testing.* Warwick & York. 1918.
Goddard, H. H. *Juvenile Delinquency.* Kegan, Paul, Trench, Trubner & Co. 3*s.* 6*d.*
Cady, V. *The Psychology and Pathology of Personality: A Summary of Test Problems.* Journal of Delinquency, VII, 225. 1922.
Cady, V. *The Estimation of Juvenile Incorrigibility.* Journal of Delinquency Monographs. 1923.
Healy, W., and Bronner, B. *Case Studies.* Judge Baker Foundation, Boston, Mass. First Series (20 studies). $2.50.
Journal of Delinquency. California Bureau of Juvenile Research, Whittier (Cal.). Bi-monthly. 30c. per number; $1.25 per annum.

APPENDIX VIII.

Examples of Tests.

(Selected for the Committee by Dr. Cyril Burt[1].)

The following extracts from tests of various representative types are reprinted here to give the reader some concrete notion of the general nature of the methods used. Any notion so gained, however, can, at the most, be very inadequate. The specimens are illustrative excerpts only. They are not intended for practical use as they stand. The material here printed has been chosen because it will be immediately intelligible to the non-psychological student, not because it includes the most efficient tests of any particular kind. As a rule, the more effective cannot be explained to those unacquainted with psychological procedure without a lengthy technical description. There is a danger, therefore, that the very simplicity of the examples here adduced may lead the reader to believe that mental testing is far more simple to carry out, and far more easy to criticise, than it actually is. In particular,

[1] Dr. Burt desires to acknowledge the help he has received from psychological and educational colleagues in the selection and description of the following tests.

tests which require special apparatus, such as the best Performance Tests, the best Vocational Tests, and the best Tests for Specific Abilities, are very imperfectly represented. The critic will find among what follows a considerable preponderance of early test-material. The earlier tests, however, are more readily understood and, having furnished models for numerous later scales, have a greater representative value.

The Committee desires to express its indebtedness to the various authors and publishers for their kindness in according permission to reprint these examples from their works.

(I) INDIVIDUAL TESTS.

(A) The Binet-Simon Scale.

(i) *Schedule of the Tests assigned to the Several Years* (*London Revision*).

The tests here enumerated include all the tests from both the 1908 and 1911 series, and are given in average order of difficulty for English children with revised age-assignments.

Age III.

1. Points to nose, eyes and mouth.
2. Repeats two numbers (one trial correct out of three), *e.g.*, 3 7, 6 4, 7 2.
3. Knows own sex, whether boy or girl.
4. Gives name and surname.
5. Names knife, key, and penny.
6. Enumerates items in two pictures out of three.

Age IV.

7. Repeats six syllables: "I am cold and hungry."
8. Repeats three numbers (one trial correct out of three): *e.g.*, 914, 286, 539.
9. Counts four pennies.
10. Points to longer of two lines.
11. Points to prettier faces in three pairs.

Age V.

12. Performs a triple order: putting key on table, shutting door, and bringing book.
13. Copies a square recognisably.
14. Repeats ten syllables
15. Gives own age.

16. Distinguishes morning and afternoon.
17. Names four primary colours: blue, yellow, green, and red.
18. Repeats four numbers.
19. Compares two weights in three pairs. differing by 9 grammes (*e.g.*, 3 and 12 grammes, 6 and 15 grammes).

AGE VI.

20. Knows (without counting) number of fingers on both hands.
21. Counts thirteen pennies.
22. Copies diamond recognisably.
23. Copies three words from script legibly, but errors allowed.
24. Names all the days of the week without error in 10 seconds.
25. Names without error four commonest coins: shilling, penny, sixpence, halfpenny.
26. Puts together two triangles of cardboard to make an oblong.
27. Defines by use three concrete words out of five (horse, chair, mother, table, fork).
28. Repeats five numbers.
29. Describes items in two pictures out of three.
30. Repeats sixteen syllables.
31. Points to right hand and left ear.

AGE VII.

32. Recognises missing features in three pictures out of four.
33. Adds without error three pennies and three halfpennies.
34. States differences between concrete objects, *e.g.*, fly and butterfly, wood and glass, paper and cardboard.
35. Writes four easy words from dictation, legibly, but errors allowed, *e.g.*, "the pretty little girls."

AGE VIII.

36. Reads prescribed passage, recalling two items out of twenty.
37. Answers two out of three easy questions, *e.g.*, "What would you do if you broke something that did not belong to you?"
38. Counts backwards from 20 to 1.
39. Gives full date.
40. Gives change for 2*d.* out of 1*s.*, actual coins being provided.
41. Repeats six numbers.

AGE IX.

42. Names the months of the year in 15 seconds.
43. Names all the nine commoner coins in 40 seconds.
44. Reads prescribed passage, recalling six items out of twenty.
45. Defines three concrete words out of five in terms superior to use (same words as for Test 27).

AGE X.

46. Arranges five small weights, differing by 3 grammes, in order of heaviness.
47. Builds two sentences with three words in 1 minute (words given, "London, money, river").
48. Copies from memory two simple geometrical designs shown simultaneously for 10 seconds.

AGE XI.

49. Explains three absurdities out of five, *e.g.*, "I have three brothers, Tom, Jack, and myself."
50. Answers three out of five harder questions, *e.g.*, "Why should we judge a person by what he does, not by what he says?"
51. Gives sixty words in 3 minutes.
52. Repeats seven numbers.
53. Builds one sentence with three words in 1 minute (same words as for Test 47).

AGE XII.

54. Gives three rhymes to the word "obey."
55. Rearranges correctly two out of three mixed sentences so as to make sense.
56. Interprets the meaning of two out of three pictures.

AGE XIII.

57. Resists suggestion given by six pairs of lines, the first three pairs differing in length.
58. Solves two circumstantial problems.

AGE XIV.

59. Repeats twenty-six syllables.
60. Defines three abstract terms (kindness, justice, charity).

AGE XV.

61. Draws a sheet of paper first folded and then cut in the middle of one edge, showing the correct shape and position of the resulting holes.
62. Gives the differences between two out of three pairs of abstract terms (idleness and laziness, poverty and misery, evolution and revolution).
63. Draws a triangle shown as part of a simple geometrical diagram, to be reversed in imagination.

AGE XVI.

64. Repeats from memory the substance of Hervieu's Aphorism on Life.
65. Gives the three chief differences between a president and a king.

(ii) *Samples of Binet's Tests (London Revision) for a Single Mental Year.*

The tests assigned to one age only are here given. The teacher should, of course, beware of inferring that these four problems are the only ones set to a child of these years.

Age VII.[1]

32.—*Recognising Missing Features.*

Materials.—Binet's four pictures of faces without mouth, nose, eye, and of a body without arms.

Procedure.—Say: "Look at this man's face. Tell me what has been left out." And, for the others, "What has been left out here?" (or "in this drawing?").

Evaluation.—Three correct answers with the four pictures are required.

33.—*Adding Three Pennies and Three Halfpennies.*

Materials.—Three pennies and three halfpennies, set out separately, but not in a row, nor all the pennies entirely apart from all the halfpennies.

Procedure.—Say: "Count this money for me; and tell me how much there is altogether."

Evaluation.—No error and no repetition of the instructions are allowed.

34.—*Stating Differences between Concrete Objects.*

Procedure.—Ask: "What is the difference between a fly and a butterfly?" ("You know what a butterfly is, don't you? And you know what a fly is? . . . They are not the same, are they? . . . In what way are they not the same?")

The following words are suggested by Binet; and the differences between them should be demanded in order:—

(i) fly, butterfly;
(ii) wood, glass;
(iii) paper, cardboard.

Evaluation.—Two out of three replies must be correct. Any true difference, though trivial, will pass. But if the child repeats the same difference for all pairs, *e.g.*, "it is larger," it is insufficient. Often a child takes a minute for one reply, but if he takes longer than 2 minutes for all he fails.

---

[1] Reprinted, by permission of the author and publisher, from Burt's *Mental and Scholastic Tests*. P. S. King & Son, pp. 43–46. (Discussion of details of procedure is here omitted.)

35.—*Writing from Dictation.*

Materials.—Pen, ink, paper.

Procedure.—Say: "Write this down for me on this piece of paper: 'The pretty little girls.'"

Evaluation.—The words must be separate, and sufficiently legible, and the spelling sufficiently accurate, for the words to be read by a person who did not know what had been dictated.

(B) A Scale of Graded Reasoning Tests.[1]

7 years.

1. Kate is cleverer than May;
   May is cleverer than Jane.

   Who is the cleverest—Jane, Kate, or May?

8 years.

2. The person who stole Brown's purse was neither dark, nor tall, nor clean-shaven.

   The only persons in the room at the time were:—

   1. Jones, who is short, dark, and clean-shaven;
   2. Smith, who is fair, short, and bearded;
   3. Grant, who is dark, tall, but not clean-shaven.

   Who stole Brown's purse?

9 years.

3. Three boys are sitting in a row.
   Harry is to the left of Willie;
   George is to the left of Harry.

   Which boy is in the middle?

10 years.

4. There are four roads here. I have come from the South and want to go to Melton; the road to the right leads somewhere else; straight ahead it leads only to a farm.

   In which direction is Melton—North, South, East, or West?

11 years.

5. Where the climate is hot, gum-trees and rubber will grow;
   Heather and grass will only grow where it is cold;
   Heather and rubber require plenty of moisture;
   Grass and gum-trees will grow only in fairly dry regions;
   Near the River Amazon it is very hot and very damp.

   Which of the above grows there?

[1] The above extracts give only one test-problem for each year, and are therefore insufficient for practical use. The complete scale is given in Burt's article on "The Development of Reasoning in School Children," *Jonrn. Exp. Pedagogy*, VI., 1921, p. 121 (with five test-problems for each year), and in an abridged form in Burt's *Handbook of Tests for Use in Schools*. P. S. King and Son, pp. 91–94.

12 *years.*

6. Field-mice devour the honey stored by the humble-bees; the honey which they store is the chief food of the humble-bees. Near towns there are far more cats than in the open country. Cats kill all kinds of mice.

Where, then, do you think there are most humble-bees—in the neighbourhood of towns or in the open country?

13 *years.*

7. A pound of meat should roast for half-an-hour;
Two pounds of meat should roast for three-quarters of an hour;
Three pounds of meat should roast for one hour;
Eight pounds of meat should roast for two hours and a quarter;
Nine pounds of meat should roast for two hours and a half.

From this can you discover a simple rule by which you can tell from the weight of a joint for how long it should roast?

14 *years.*

8. John said: "I heard my clock strike yesterday, ten minutes before the first gun fired. I did not count the strokes; but I am sure it struck more than once, and I think it struck an odd number."

John was out all the morning from the earliest hours, and his clock stopped at five to five the same afternoon.

When do you think the first gun fired?

(C) Performance Tests.

1. *Healy Picture Completion, Test II.* (United States Army Scale.

This test consists of ten parts, each representing a scene in a school-boy's day—for example, he is shown eating breakfast in one part, on his way to school in the next, in his class in a third, and so on. In each part some essential thing is missing and has to be filled in with insets selected from a number by the subject. This test is based on typically American scenes, and may, therefore, in some ways be unfair to English children since the material is less familiar to them.

(Summarised from the Official Report as contained in Vol. 15 of the *Memoirs of the National Academy of Science.*)

2. *Cube Construction Test.* (United States Army Scale.)

This test consists of three models. Model 1 is a block of wood, 1 by 3 by 3 inches, painted a dark red on the four sides—not on the upper and lower surfaces—and cut to a depth of 2 mm., so that it closely resembles a composite of nine small cubes. Nine 1-inch cubes necessary for the construction of model 1 make up the material for the first part of this test. Of these cubes four are painted red on two sides, four on one side, and one is not painted at all. For model 2 there is a block of wood the same size as model 1, but painted on the top as well as the four sides and bottom. There are also nine 1-inch cubes necessary for the construction of model 2. A 2-inch cube (model 3) is unpainted and cut on the six surfaces so that it looks like a composite of eight small cubes. There are also eight 1-inch cubes painted on three sides only for the construction of model 3. The examiner shows model 1, and the small cubes for its construction, pointing out the painted and unpainted sides of both model and cubes. He then puts the cubes together with the minimum number of placings, after which he presents the small cubes in an irregular order to the subject, who is to put them together. The same procedure is followed with models 2 and 3, except that the examiner does not put the blocks together first as in model 1. The scoring of this test is based on the time taken and the number of moves or placings of the cubes. Each model is scored separately and the three scores added together for the total score.

(Summarised from the Official Report as contained in Vol. 15 of the *Memoirs of the National Academy of Sciences.*)

3. *Cube Imitation Test.* (Pintner & Paterson.[1])

In this test, four black 1-inch cubes are placed in a row before the subject. With the fifth cube the examiner taps the four in a given order. The subject then imitates this scheme of taps by tapping the cubes in the same order. There are twelve such patterns of increasing difficulty. The scoring is based on the number of patterns correctly imitated by the subject.

4, 5 & 6. *Triangle and Diagonal Tests, and Healy Puzzle "A."* (Pintner & Paterson.[1])

These are small formboards, all measuring about 17 by 12.8 by 1 cm. In the first test there are two recesses—one a rectangle and the other an equilateral triangle—which are to be filled by the subject with four right-angled triangles. In the diagonal test there is one large rectangular recess, which is to be filled with two right-angled triangles, one smaller right-angled triangle, one rectangle and

[1] Pintner and Paterson, *A Scale of Performance Tests.* D. Appleton & Co., New York, 1917. pp. 67, 69, and pp. 40–53 (Quoted by permission of the authors, and of Messrs. D. Appleton & Co., of New York, the owners of the copyright.)

one quadrilateral. Healy Puzzle "A" also has one large rectangular recess only, which is to be filled with five rectangles, two of which are the same size. Scoring in these three tests is based on time taken for a correct solution. No credit is given for partial or incorrect solutions.

(II) GROUP TESTS.

The following are illustrative samples, first, of some of the earliest group-tests—Opposites, Analogies, and Completion—used in this country and elsewhere, and recurring again and again in later compilations, and, secondly, of the later compilation known as the American Army Tests, one of the first compilations or "batteries" (as they are termed in America) to be used upon an extensive scale. Recent English compilations are represented by the Northumberland and the Chelsea collections.

(A) Opposites.

Instructions. Against each printed word write down, in the space left, a word which is the opposite of the printed word. (The instructions are usually accompanied by examples explained upon the blackboard.)

	Answer.		*Answer.*
1. Old	.....	26. Warm	.....
2. Poor	.....	27. Tight	.....
3. Big	.....	28. Mountain	.....
4. Early	.....	29. Father	.....
5. Long	.....	30. True	.....
6. Easy	.....	31. Shut	.....
7. Inside	.....	32. Female	.....
8. Pretty	.....	33. Few	.....
9. Boy	.....	34. Heavy	.....
10. Wet	.....	35. Multiply	.....
11. Kind	.....	36. Absent	.....
12. Winter	.....	37. Moving	.....
13. Woman	.....	38. Question	.....
14. Slow	.....	39. Now	.....
15. White	.....	40. Polite	.....
16. Upwards	.....	41. East	.....
17. Loud	.....	42. Enemy	.....
18. Crooked	.....	43. Nobody	.....
19. Cheap	.....	44. Glad	.....
20. Busy	.....	45. Top	.....
21. Sunrise	.....	46. Possible	.....
22. Brother	.....	47. Come	.....
23. Borrow	.....	48. Front	.....
24. Clean	.....	49. Day	.....
25. Common	.....	50. Tame	.....

With tests of the foregoing and similar types, to secure uniformity of marking, alternative answers, one right and two or three wrong, are sometimes presented to the candidate. When, as in the foregoing example, the candidate is left to discover and write his own answer, different answers of varying adequacy are often given: *e.g.*, as the opposite of "*Common*"—"rare," "uncommon," "infrequent," "seldom," or "not common" may be suggested; and different examiners might mark these differently. The usual device for overcoming this difficulty is illustrated in the following tests.

(B) ANALOGIES.[1]

(Time allowed, 3 minutes.)

Instructions. In each question a fourth word is wanted which goes with the third word (in capitals) in the same way as the second word (in capitals) goes with the first. Look in the second line of each question for the word that is wanted; and draw a line under it. Do not write anything.

Examples.

GOOD is to BAD as WHITE is to
CLEAN, BLACK, WICKED, RED.

BAKER is to BREAD as TAILOR is to
TAILORESS, CAKE, MAN, CLOTHES.

1. FATHER is to MOTHER as HUSBAND is to
RED, WIFE, GREEN, BUSINESS.
2. UP is to DOWN as HIGH is to
LOW, BOOK, COAL, DIFFICULTY.
3. PRINCE is to PRINCESS as KING is to
DUCHESS, CROWN, QUEEN, ROYAL.
4. PARENT is to CHILD as MOTHER is to
WIFE, MAID, DAUGHTER, SERVANT.
5. FIRE is to HOT as ICE is to
CREAM, WATER, SOLID, COLD.
6. EAT is to BREAD as DRINK is to
DRUNKARD, THROAT, CUP, WATER.
7. SITTING is to CHAIR as SLEEPING is to
WALKING, TIRED, BED, DREAM.
8. JANUARY is to DECEMBER as SUNDAY is to
TUESDAY, MONDAY, SATURDAY, WINTER.
9. FLYING is to BIRD as CREEPING is to
AEROPLANE, SNAIL, GROUND, FLOWER.
10. TEARS are to SORROW as LAUGHTER is to
JOY, SMILING, CRYING, MISERY.

[1] This and the following series of text-problems are from collections of group-tests prepared for the National Institute of Industrial Psychology.

11. SIGHT is to PICTURE as HEARING is to
SONG, COLOUR, EAR, SEEING.
12. EGG is to BIRD as SEED is to
PLOUGHMAN, FOWL, PLANT, WHEAT.
13. REMEMBER is to PAST as ANTICIPATE is to
FANCY, FUTURE, FORGET, PRESENT.
14. BEAR is to CUB as DOG is to
CAT, SPANIEL, PUPPY, KITTEN.
15. FACT is to FICTION as HISTORIAN is to
HISTORY, BOOK, NOVELIST, MATHEMATICIAN.
16. BEAUTY is to ART as TRUTH is to
SCIENCE, MUSIC, ARTIST, LIAR.
17. ASLEEP is to AWAKE as DEAD is to
HEAD, CORPSE, ALIVE, MORTALITY.
18. FOOD is to MAN as FUEL is to
WOMAN, STEAM, ENGINE, VAPOUR.
19. SKY is to GROUND as CEILING is to
GAS, WALL, FLOOR, CHANDELIER.
20. SWEET is to HONEY as SOUR is to
SUGAR, SALT, VINEGAR, PEPPER.
21. HORSE is to MULE as DOCILE is to
RIDER, STUBBORN, DONKEY, MAN.
22. WHEN is to WHERE as TIME is to
HOW, WHY, SPACE, LENGTH.
23. MOTIVE is to METHOD as WHY is to
WHERE, MANNER, REASON, HOW.
24. CAUSE is to EFFECT as DISEASE is to
REASON, CONSEQUENCE, DEATH, LIFE.
25. THE DAY BEFORE YESTERDAY is to THE DAY AFTER TO-MORROW as SATURDAY is to
SUNDAY, MONDAY, WEDNESDAY, FRIDAY.

(C) Completion Test.

(Seven problems only are given from a test-sheet of twenty-five.)

Instructions. Underline the word, phrase, or number that makes the best sense, wherever there are three printed one above the other.

1. Grass is wet, green, but the sky is green, wet. blue, blue.

2. The man rode fell, climbed off his bicycle and broke, changed his arm. cured,

3. I saw a beggar in the hotel, street, parlour and showed him sixpence to sell, owed, gave steal, buy some food.

4. If Robinson runs faster than Jones, and Jones runs faster
Brown / Smith
than Brown, then Smith runs slowest / fastest / backwards of the three.

5. If to-day were
Friday, / Saturday, / Sunday,
then the day
before / after / preceding
to-morrow
would be
Thursday. / Tuesday. / Wednesday.

6. It has been argued that Mahomed was
once / both / either
an enthusiast / ignoramus; / imposter; / evangelist;
and, were this true, those who / may / will deny that
Mahomedannism / Christianity / he
was an
infidel / insincerity / enthusiast
would be forced to conclude that
he must / he should / could not
have been an imposter.

7. A man, writing on 1st January, 1922, said: "My sister, who
died / was born / was married
on 13th November,
1898, / 1858, / 1900,
will be
twenty-five / thirty-three / thirty-four
years old
next November." / year." / month."

(D) American Army Tests. ("Group Examination Alpha.")[1].

The following extracts show the type of problem used in each of the eight tests, and the general nature of the instructions. This scheme of tests has formed a model for innumerable sets of group-tests. As illustrated by this early experiment, many defects in the test-material and in the general form of the examination, will be obvious to the reader, many of which have been variously amended in subsequent forms of such tests.

[1] Volume 15 of the *Memoirs of the National Academy of Sciences*, Washington, U.S.A.

Test 1.

The first page of the candidate's booklet contained rubrics and diagrams, as illustrated below.

The headings having been filled, the examiner commences with the following instructions :—

1. "'Attention'! 'Attention' always means 'Pencils up.' Look at the circles at 1. When I say 'go,' but not before, make a cross in the first circle and also a figure 1 in the third circle—Go!" (Allow not over 5 seconds.)

4. "'Attention'! Look at 4. When I say 'go' make a figure 1 in the space which is in the circle, but not in the triangle or square, and also make a figure 2 in the space which is in the triangle and circle, but not in the square—Go!" (Allow not over 10 seconds.)

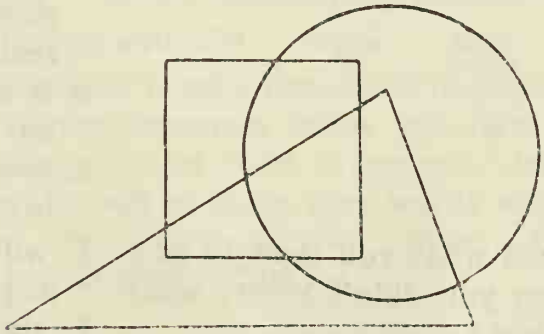
(N.B. *Examiner.* In reading 5, don't pause at the word CIRCLE as if ending a sentence.)

5. "'Attention'! Look at 5. If a machine gun can shoot more bullets a minute than a rifle, then (when I say 'go') put a cross in the second circle; if not, draw a line under the word No—Go!" (Allow not over 10 seconds.)

FORM 5. GROUP EXAMINATION ALPHA. GROUP NO.....
Name..... Rank..... Age.....
Company..... Regiment.....Arm..... Division.....
In what country or state born?.....Years in U.S.?.....
Race.....Occupation.....Weekly Wages.....
Schooling: Grades 1, 2, 3, 4, 5, 6, 7, 8: High or Prep. School, Year 1, 2, 3, 4. College, Year 1, 2, 3, 4.

1.

4.



5.

Yes. No.

Test 2.

Instructions. Get the answers to these examples as quickly as you can. Use the side of this page to figure on if you need to.

Samples.

1. How many are five men and ten men? .. Answer (15)
2. If you walk 4 miles an hour for 3 hours, how far do you walk? .. .. Answer (12)
1. How many are thirty men and seven men? .. Answer ( )
4. Mike had twelve cigars. He bought three more, and then smoked six. How many cigars did he have left? .. .. Answer ( )
5. A company advanced 5 miles and retreated 3 miles. How far was it then from its first position? .. .. Answer ( )

Test 3.

Instructions. This is a test of common sense. Below are sixteen questions. Three answers are given to each question. You are to look at the answers carefully; then make a cross in the square before the best answer to each question, as in the sample :—

Sample.

Why do we use stoves? Because .. .. they look well.
they keep us warm.
they are black.

Here the second answer is the best one and is marked with a cross.

Begin with No. 1 and keep on until time is called.

1. Cats are useful animals, because .. .. they catch mice.
they are gentle.
they are afraid of dogs.

4. Why judge a man by what he does rather than by what he says? Because .. what a man does shows what he really is.
it is wrong to tell a lie.
a deaf man cannot hear what is said.

5. If you were asked what you thought of a person whom you didn't know, what should you say? .. .. .. I will go and get acquainted.
I think he is all right.
I don't know him and can't say.

Test 6.

Instructions. Look at each row of numbers below, and on the two dotted lines write the two numbers that should come next.

Samples.

	2	4	6	8	10	12	..14..	..16..
	9	8	7	6	5	4	...3..	...2..
	2	2	3	3	4	4	...5..	...5..
	1	7	2	7	3	7	...4..	...7..
1.	3	4	5	6	7	8	.....	.....
2.	10	15	20	25	30	35	.....	.....
3.	8	7	6	5	4	3	.....	.....
4.	3	6	9	12	15	18	.....	.....
5.	5	9	13	17	21	25	.....	.....

Test 7.

Instructions. In each of the lines below, the first two words are related to each other in some way. What you are to do in each line is to see what the relation is between the first two words, and underline the word in heavy type that is related in the same way to the third word. Begin with No. 1 and mark as many sets as you can before time is called.

Samples.

Sky — blue : : grass —table green warm big.
Fish — swims : : man —paper time walks girl.
Day — night : : white — red black clear pure.

1. Gun — shoots : : knife — RUN CUTS HAT BIRD.
2. Ear — hear : : eye — TABLE HAND SEE PLAY.
3. Dress — woman : : feathers — BIRD NECK FEET BILL.
4. Handle — hammer : : knob — KEY ROOM SHUT DOOR.
5. Shoe — foot : : hat — COAT NOSE HEAD COLLAR.

Test 8.

Instructions. Notice the sample sentence :

People hear with the eyes ears nose mouth.

The correct word is ears, because it makes the truest sentence.

In each of the sentences below you have four choices for the last word. Only one of them is correct. In each sentence draw a line under the one of these four words which makes the truest sentence. If you cannot be sure, guess. The two samples are already marked as they should be.

Samples.

People hear with the eyes ears nose mouth.
France is in Europe Asia Africa Australia.

1. America was discovered by Drake Hudson Columbus Balboa.
2. Pinochle is played with rackets cards pins dice.
3. The most prominent industry of Detroit is automobiles brewing flour packing.
4. The Wyandotte is a kind of horse fowl cattle granite.
5. The U.S. School for Army Officers is at Annapolis West Point New Haven Ithaca.

(E) The Northumberland Mental Tests.[1]

These tests were six in number, arranged in a small booklet. The first part of the booklet contained detailed instructions and sample exercises, and the latter part the tests proper. Examples of each are given below.

Test B1.

Cross out plainly the " extra " word in each of the following lines :—

charity	kindness	benevolence	revenge	love.
square	circular	oblong	hexagonal	triangular .
needle	tack	nail	knife	pin.
coal	bread	coke	wood	paper.
bran	wool	cotton	hemp	jute.
hair	feathers	wool	grass	fur.

Test B2.

Give the number that comes next in each of the following lines :—

1	3	9	27	81	243	( )
17	15	13	11	9	7	( )
1	2	4	7	11	16	( )
96	48	24	12	6	3	( )
7	5	5	7	5	5	( )
9	8	7	6	7	8	( )

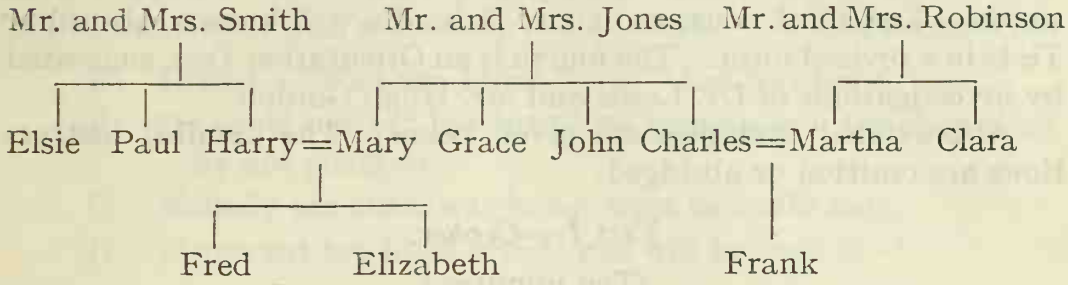
Test B3.

Cross out the " extra " number in each of the following lines :—

26	3	7	31	13	17
18	22	30	24	6	12
81	27	11	1	9	3
8	2	16	32	6	4
7	28	21	35	14	27

[1] Professor Godfrey Thomson, *British Journal of Psychology*, Vol. XII, Part 3, pages 201 et seq. Reprinted by permission of Professor G. Thomson, and Mr. F. C. Bartlett, Editor of the *British Journal of Psychology*, and Messrs. Harrap & Co., Ltd , the present publishers.

Test B4.



From the above diagram answer the following questions :—

	Answers.
What is Frank's surname ?	.....
What relation is Elizabeth to John ?	.....
What relation is Elizabeth to Frank ?	.....
What relation is Elizabeth to old Mr. Jones ? ..	.....
How many aunts has Frank ?	.....
How many uncles has Fred ?	.....
Who are Frank's grandparents ?	.....

Test B5.—*Part* 1.

In your mind (without writing them down) you have to arrange the five words in each line below in the proper order and then underline the middle word of this order :

paragraph	volume	chapter	sentence	letter.
house	street	room	town	county
orange	red-currant	grape	plum	melon.
general	sergeant	captain	private	corporal.

Test 6.

In each foreign sentence underline the word which corresponds to the underlined word in the English sentence :

Mausam badalta hai.	It is changeable weather.
Achchha khub mausam hai.	It is good weather.
Sirf das baje hai.	It is only ten o'clock.
Das gaz napo.	Measure ten yards.
Ham kahan jaen ?	Where shall we go ?
Ham chha baje khama khawenge.	We shall dine at six o'clock.
Yih rah kahan jati hai ?	Where does this road lead to ?

(F) The Chelsea Mental Tests.[1]

This set contains four tests, first used by Dr. Ballard in the schools of the Chelsea Division of London. The first test is a Cipher Test ; and is the only one for which a time-limit is applied. The

[1] Reprinted, by permission of the author and publisher, from P. B. Ballard's, *Group Tests of Intelligence*, Hodder and Stoughton. pp. 117–128

second is intended to gauge the pupils' power of understanding words. The third consists of Dr. Ballard's well-known Absurdity Tests in a revised form. The fourth is an Orientation Test, suggested by investigations of Dr. Lewis and Mr. Hugh Gordon.

Abbreviated examples are given below. The detailed instructions are omitted or abridged.

Test I.—Cipher.

(Ten minutes.)

Key :	a	e	i	o	u	h
	,	.	;	:	!	-

Answer these questions :—

1. C,n p;gs fly ?
2. W-,t c:l:!r ;s gr,ss ?

25. ;f , m,n st,rts d;gg;ng ,t n::n w-,t t;m. w;ll ;t b. w-.n -. -,s w:rk.d f:r f;v. -:!rs ?

Test II.—Meanings of Words.

Look at this sentence : World, football, marble, melon are all (solid, eatable, round, small). Only one of the four words in brackets would make the sentence true—"round." Answer the first eight questions by writing down for each one the word that makes the sentence true.

1. Honey, jam, saccharine, treacle are all (liquid, sweet, sticky, manufactured).

8. Birthdays, pain, error, death are all (disagreeable, inevitable, interchangeable, unconquerable).

Now work the rest of the paper.

22. When the moon grows larger it waxes, when it gets smaller it ————. What is the missing word ?

25. Which of these words means the largest number of things : cabs, barrows, trams, vehicles, carriages, omnibuses, carts, motor-cars ?

Test III.—Absurdities.

After each statement there are four tries (A, B, C, D) at saying what is foolish in it. Find out which of the four is the best.

1. A boy who wanted to go to a cinema, but had no money, thought it would be a good plan to walk in backwards, for the man at the door would think he was going out, and would not ask for his ticket. Foolish because :—

A. You cannot go in without a ticket.
B. The boy was trying to cheat.
C. The door-keeper would see that he was moving inwards.
D. Cinemas are bad for the eyes.

2. A countryman came up to London during the war, and saw on a hoarding the words: "Eat less bread; do it now." He immediately went into a tea-shop to do it now. Foolish because :—

E. It said: "Eat less bread," not "Eat less buns."

F. He could not eat less bread by going into a tea-shop than by not going in.

G. Nobody ate more war bread than he could help.

H. If you eat too little bread, you will become ill.

24. While standing near a clock-tower just before the clock struck twelve two boys tried to find out which of them could hold his breath the longer. Neither of them won, for one was able to hold his breath from the first stroke of the clock to the sixth, and the other from the sixth to the twelfth. Foolish because :—

W. The boys could not hold their breath so long.

X. It is unhealthy to stop breathing.

Y. They ought to have started together.

Z. The second boy won.

Test IV.—Orientation.

1. Draw the capital letter F as it would appear if it were printed upside down.

3. If a boy stands on his head with his face to the south, where will his right hand point ?

4. If these six letters were seen reflected in a mirror, some would look the same as they do here and some would look different. Write down those that would look the same : S P M V N H.

Name by letters all the people you see drawn here whose left leg is nearer the beginning of the line (nearer the X) than the right leg.

(Six drawings follow showing a boy in various positions, walking to the right or to the left, towards the reader or away from him, and standing on his head facing the reader or with his back to him.)

(IV) Tests of Specific Abilities.

The testing of specific abilities generally requires laboratory apparatus and a laboratory technique such as does not lend itself to simple description. The following, however, is one of the simplest tests available.

Test of Mechanical Memory.[1]

Material. The following lists of concrete and abstract monosyllables :—

Three-term lists.

Concrete.	*Abstract.*
Street	Time
Ink	Art
Lamp	Route

Four-term lists.

Concrete.	*Abstract.*
Spoon	Phase
Horse	Work
Chair	Truth
Stone	Thing

Five-term lists.

Concrete.	*Abstract.*
Ground	Tact
Pen	Scope
Clock	Proof
Boy	Scheme
Chalk	Form

Six-term lists.

Concrete.	*Abstract.*
Desk	Space
Milk	Creed
Hand	Pride
Card	Guile
Floor	Pledge
Cat	Cue

Seven-term lists.

Concrete.	*Abstract.*
Ball	Craft
Sponge	Myth
Glass	Rate
Hat	Cause
Fork	Style
Stove	Youth
Post	Mood

Eight-term lists.

Concrete.	*Abstract.*
Coat	Law
Girl	Thought
House	Plot
Salt	Glee
Glove	Life
Watch	Rhythm
Box	Faith
Mat	Mirth

To test sheer rote memory meaningless "nonsense syllables" are frequently used—*e.g.*, "jad, mig, dep"; to test logical memory, words suggesting obvious connections—*e.g.*, "grass, green, blue, sky, star, moon, sun."

Method. For group-tests the lists are simply recited once to the children, the words being enunciated at the rate of one per second. Before each list is given, the examinees are notified of the number of words to be given. After each list has been read, the subject writes down as many words as he can remember upon prepared blanks.

For individual tests it is best to allow the subject to see the words one by one, and one at a time, by means of some simple exposure-apparatus. Mechanical appliances, working accurately to time, are generally used; but a sheet of cardboard with a slot in it

[1] Reprinted, by permission of the author and publishers, from Whipple's *Manual of Mental and Physical Tests.* Warwick and York. Baltimore, 1910. Test 38. pp. 356 et seq.

rather larger than the words may be used, moved over the page so as to screen all but the word shown. The experimenter reads the words as each is shown, and the child pronounces them with him, so that he simultaneously sees, hears, and utters each list.

Evaluation. The simplest method of marking is to count up the number of words correctly reproduced from memory. Various methods of classifying errors have been suggested, but are not essential unless a detailed psychological analysis is to be undertaken. If the lists are redictated and again reproduced after a larger interval—of a day, week or fortnight—some relative measure of " long-distance memory " is obtained.

(V) Standardised Scholastic Tests[1].

The following examples are selected to illustrate how the principle of an age-scale may be applied to the testing of scholastic attainments. The first test—for mechanical reading—is given in full. The figures at the side indicate how the results are marked ; a child who reads thirty-five words out of the whole list—*e.g.*, down to " return " but no further—scores a " mental age " for mechanical reading of 7.5 years.

(i) *Reading (Graded Vocabulary Test).*

Age last Birthday.					
4.	to	is	of	at	he
	my	up	or	no	an
5.	his	for	sun	big	day
	sad	pot	wet	one	now
6.	that	girl	went	boys	some
	just	told	love	water	things
7.	carry	village	nurse	quickly	return
	known	journey	terror	obtain	tongue
8.	shelves	scramble	twisted	beware	commenced
	scarcely	belief	steadiness	labourers	serious
9.	projecting	fringe	luncheon	nourishment	overwhelmed
	urge	explorer	trudging	events	motionless
10.	economy	formulate	exhausted	contemptuous	
	universal	circumstances	destiny	glycerine	
	renown				
	atmosphere				

[1] Reprinted, by permission of the author and publishers, from Burt's *Handbook of Tests for use in Schools.* P. S. King and Son. pp. 2, 3, 16, 19, et seq.

11.	perpetual	emergency	humanity	perambulating
	apprehend	excessively	domineer	theory
	ultimate			
	reputation			
12.	physician	fatigue	philosopher	melodrama
	constitutionally	champagne	encyclopedia	hypocritical
	autobiography			
	efficiency			
13.	melancholy	exorbitant	influential	terminology
	mercenary	contagion	fallacious	binocular
	palpable			
	microscopical			
14.	atrocious	phlegmatic	refrigerator	unique alienate
	eccentricity	ingratiating	subtlety	poignancy phthisis

With two following tests the words and problems for *one age only* are here extracted. Ten words or ten sums alone are, of course, insufficient to determine the mental age.

(ii) *Spelling* (*Graded Vocabulary Test*).

Age 7.

table	even	fill	black	only
coming	sorry	done	lesson	smoke

(iii) *Arithmetic* (*Graded Oral Test : Mental*).

Age 7.

1. My brother has picked six nuts, my sister has picked ten, and I have picked eighteen. How many have we got altogether ?

2. Twelve girls have a farthing each. How many pennies is that ?

3. How many $\frac{1}{2}d.$ stamps can I buy for $9d.$?

4. I started with fourteen marbles, and I have won twenty-six. How many have I now ?

5. I have $2s.$ to divide among four children. How much should each have if all are to have the same amount ?

6. How many days are there in six weeks ?

7. My brother is 4 feet high. How many inches is that ?

8. On a tram there were fifty people who each paid $1d.$ fare. How much (in shillings and pence) did the conductor take altogether ?

9. If treacle were $8d.$ a pound, how much would $\frac{3}{4}$ lb. cost ?

10. Yesterday we went blackberrying. I picked twenty-one berries, and my brother ate twelve of them. How many were left ?

(VI) Vocational Tests.

The following scheme of tests, devised at the National Institute of Industrial Psychology, illustrates the method of testing for two particular occupations.

(A) *Tests for Shorthand-Typists*[1].

I. Graded Tests of Intelligence :—
- (1) Discriminating Opposites and Synonyms (alternative answers given).
- (2) Analogies (alternative answers given, one correct and three incorrect for each problem).
- (3) Mixed Sentences (the examinee being required to indicate the truth or falsity of the rearranged sentences).
- (4) Completion of Sentences (the examinee being required to find the missing words).

II. Graded Tests of Educational Attainments :—
- (1) Spelling.
- (2) Arithmetic.

III. Tests of Linguistic Ability and General Information :—
- (1) Synonyms. (The examinee is required to think of and write down a word of similar meaning for each of those given. The given words are chosen to throw light not so much upon intelligence as upon range of vocabulary and literary precision, and incidentally upon certain temperamental and moral qualifications.)
- (2) Definitions. (The examinee is required to give the meaning of simple technical words. The words are especially chosen to test familiarity with some twenty-five fields of knowledge, one easy and one hard term being chosen for each.)

The foregoing are group tests.

IV. Tests of Typewriting and Shorthand Writing :—

A. Shorthand.

- (1) Graded Speed Tests. These consist of a specially prepared series of six passages containing 60, 80, 100, 120, 140, and 160 words. The passages are arranged so that each is divided into twelve equal sections; and each section is to be read within the space of 5 seconds, so that the entire passage occupied 1 minute. This device makes it possible for the experimenter, with the aid of a watch showing seconds, to read the passage at a uniform speed and at the

[1] From "Tests for Clerical Occupations," *Journ. Nat. Inst. Ind. Psychol.* Vol I, No. i. pp. 23–27. (Quoted by permission of the Council of the National Institute of Industrial Psychology.)

specific rate. The candidates transcribe their shorthand notes back into longhand; and from the proportion of correct words in each of the passages it is possible to compute, by means of an arbitrary formula, the approximate speed at which the individual could take down dictated matter in shorthand.

(2) Outline Tests. This consists of a series of fifty somewhat unusual words graded in difficulty and exemplifying the more important principles of Pitman's shorthand. There is no time-limit to this test. The candidates transcribe their outlines back into longhand; and the results are marked both for accuracy of outline and correctness of transcript (disregarding spelling).

B. Typewriting.

(1) Speed Tests.

(i) Typing from Memory (with time limit). The candidate is required to type over and over again for two minutes a short sentence known by heart, such as "Mary had a little lamb."

(ii) Typewriting from copy (with amount limit). The candidate is given a printed passage to copy at maximum speed on the typewriter; and the time is recorded with a stop-watch.

(2) Tests of Accuracy and Display.

(i) Accuracy. The candidate is given a manuscript containing a large number of errors and of corrections inserted after the fashion of a corrected printer's proof; and is required carefully to type from this a fair copy. The result is marked both for speed and for accuracy.

(ii) Display. A long announcement of a meeting, typed without capitals or punctuation or any form of arrangement or indentation, is given to the candidate; and the candidate is required to retype the matter, arranging it as suitably as possible. The result is marked for excellence of display.

(iii) Tabulating. A complicated list of data typed consecutively is given to the candidate, who is required to retype and rearrange it in tabular form, adding certain simple calculations.

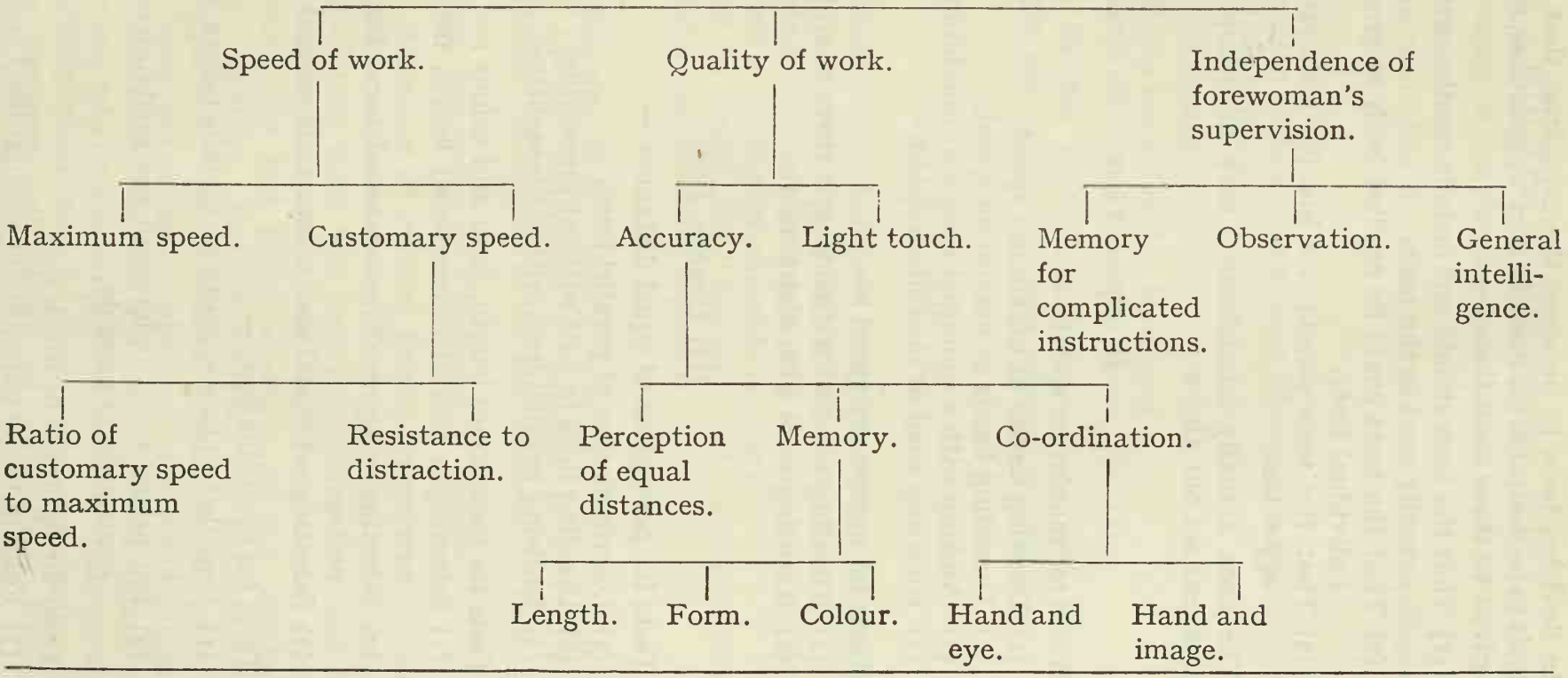
(3) Manuscript Reading. Two letters are given to the candidate, written in a very illegible hand; the candidate is required to re-transcribe them. The test is marked for accuracy and speed.

(B) *Tests for Dressmakers' Apprentices*[1].

The foregoing tests for shorthand-typists are largely tests of acquirement; the following tests for dressmakers' apprentices are intended primarily to measure natural aptitude.

The analysis of the occupational requirements, upon which the tests are based, is shown in the following scheme:

PSYCHOLOGICAL FACTORS DETERMINING EFFICIENCY OF DRESSMAKERS' APPRENTICES.



[1] From "Vocational Tests for Dressmakers Apprentices," by Winifred Spielman. *Journ. Nat. Inst. Ind. Psychol.*, I, vii. 1923. pp. 277 et seq. (Quoted by permission of the Council of the National Institute of Industrial Psychology).

In devising tests to measure the capacities that were found by analysis to be desirable for dressmakers' apprentices, special attention was given to three conditions :—

(1) That the tests should not include needlework or any process actually used in the trade.
(2) That the tests could be applied both as group tests and as individual tests.
(3) That the tests should not involve elaborate or expensive apparatus.

The tests actually administered may be grouped according to the analysis set out above.

(I) *Speed Tests.*

(*a*) Tests for maximum speed :—

(1) Drawing circles at maximum speed.
(2) Threading beads at maximum speed.
(3) Pricking with a mounted needle at maximum speed.
(4) Knotting wool at maximum speed.

(*b*) Tests for customary speed :—

(1) Knotting wool with direction, " there is no hurry."
(2) Knotting wool with distraction.

(II) *Quality Tests.*

(*a*) Tests for perception of equal distances :—

(1) Discrimination of parallel lines.
(2) Dividing lines in half without measuring.
(3) Drawing parallel lines without measuring.

(*b*) Tests for memory of length, form and colour :—

(1) Selecting a line of remembered length from several alternatives.
(2) Selecting a figure of remembered form from several alternatives.
(3) Selecting a remembered colour from several alternatives.

(*c*) Tests for co-ordination :—

(1) Pencilled dots are made at certain points as accurately as possible.
(2) Pin pricks from underneath are made at certain points as accurately as possible.

(*d*) Tests for lightness of touch :—

(1) Folding tissue paper under standardised conditions.

(III) *Tests of Independence.*

(*a*) Tests for memory of complicated instructions :—

(1) " Directions " test.

(2) Reproduction after an interval of a read description of a dress.

(*b*) Tests for observation :—

(1) Enumeration of details of a dress (shown in a fashion plate) directly it is removed.

(*c*) Tests of general intelligence :—

(1) Binet-Simon tests (London Revision).

(2) Group tests (National Institute Series, No. 33).

In addition to the psychological tests mentioned above, a standard eye-sight test is given in cases where it is desirable.

APPENDIX IX.

The views of various psychologists who gave oral evidence before the Committee or sent memoranda on the following points :—

(*a*) The factors involved in " general " ability which must be included in testing for educable capacity.

(*b*) The need for testing for " special " and " group " abilities in determining educable capacity.

Mr. William Brown, M.D., D.Sc., M.R.C.P., *late* Reader in Psychology in the University of London (King's College), Reader in Mental Philosophy in the University of Oxford.

Factors Involved in " General " Ability.

There was a difference of opinion between psychologists as to whether or not there was one central factor known as " general " ability. Some said that there was such a factor, others said that there was not, whilst others again held that there were " group " abilities. He personally had found evidence of the correlation between " group " abilities. But the evidence was not sufficiently convincing to support the view that such correlation was governed by one central factor pervading all abilities. He was inclined to favour the view that there was such a thing as general ability with sub-factors. At any rate he was of opinion that the factors involved in " general " ability could only be determined empirically, with the aid of the mathematical theory of correlation. Professor Spearman, working from the standpoint of the mathematical theory

of correlation regarded his results as evidence of one central factor. Professor Thomson said that if one took a number of chance abilities and worked out the results one would get by chance the same hierarchy. Professor Thomson said, therefore, that the hierarchy was due to chance, but witness was not convinced that that view was right. He did not consider that such a divergence of views would affect the use of mental tests.

There was evidence in favour of the existence of both "special" and "group" abilities, and, therefore, both classes of abilities should be tested for in determining educable capacity.

Mr. CYRIL BURT, D.Sc., Psychologist to the London County Council.

FACTORS INVOLVED IN "GENERAL ABILITY."

Dr. Burt stated that the educational capacity of a child at any period of his life might be assumed to depend on mental factors of two kinds: (i) inborn psychological abilities of a relatively elementary and general nature; and (ii) acquired capacities of a more complex and specific character, chiefly memories and habits, such as particular items of knowledge and particular forms of skill. The distinction between the inborn and acquired capacities was mainly theoretical.

Of inborn psychological capacities, "general intelligence" was by far the most important, as it was also by far the most easily tested. "General intelligence" comprised a number of factors, but witness would regard it more as a single complex quality than as a group of independent elements. It was best measured by tasks requiring the voluntary maintenance of attention, quick and accurate learning (in the broader sense of the word, namely, adaptation to relatively novel conditions), and on the higher mental levels, reasoning. These should, perhaps, be regarded rather as modes of general ability, than as elements entering into general ability as component factors. Of other inborn elementary capacities, the next most important for educational purposes was probably that which might be described as "long-distance mechanical memory." By "long-distance" memory witness meant a memory that could retain its powers over a period of weeks or months rather than of minutes; and by "mechanical memory" he meant recollection of things in arbitrary connection, *e.g.*, historical dates rather than the gist of a continuous passage of prose. The factors he had mentioned were not the only factors, but probably the most important. There was, for example, another factor of lesser importance, namely, imagery. These inborn abilities were inherited at birth and could

not be inculcated. There was, therefore, a point of maximum attainment fixed by congenital limitations beyond which a child could not go, and to which many in fact did not reach.

Testing of "Special" and "Group" Abilities.

Like most intellectual characteristics that had been subjected to statistical analysis, educational attainments appeared to depend upon capacities of two orders: (i) a more general capacity entering into all subjects alike—a capacity in which "general intelligence" was the chief factor; (ii) special capacities limited to particular subjects or particular groups of subjects. Of these latter, the most important, or at least the most easily demonstrable in the ordinary elementary school, appeared to be: (*a*) linguistic or literary ability; (*b*) arithmetical ability; and (*c*) manual ability, to which might perhaps be added (*d*) artistic capacity; and (*e*) musical capacity. Dr. Burt considered that it was possible, and would be useful, to develop a series of tests for measuring most of these capacities in schools.

Mr. James Drever, D.Phil., Combe Lecturer in Psychology at the University of Edinburgh.

General Ability.

"General ability" is a somewhat indefinite expression. Its meaning will necessarily vary according to the criterion with reference to which ability is determined. The acid test of life may lead to a verdict very different from that based upon school or academic success, and in precisely the same way school success may not bear out the verdict based upon the results of our special tests for "general ability." Personally I think I succeed in keeping matters clear in my own mind by distinguishing between "tendency" and "capacity" in the natural endowment of the human being. In discussing the subject of mental testing, it is certain that some such distinction must be drawn and adhered to. As a general rule mental tests are tests of "capacity" and do not profess to be more. Indirectly and incidentally they may throw valuable light on "tendency," but they are not specifically adapted for this purpose, and it would seem well nigh impossible to adapt them, at least on any basis of mental testing yet proposed. In what follows it must therefore be understood that "general ability" and "educable capacity," though really including factors of the "tendency" order —especially the former—are, for the purpose of the present discussion, being taken more narrowly, and in the strict sense of "capacity" as opposed to "tendency."

Miss Beatrice Edgell, Ph.D., Lecturer in Philosophy, Bedford College, and University Reader in Psychology in the University of London.

Factors Involved in General Ability.

The factors to be included fell into three groups :

(I) The processes basic to, and constitutive of, intellect, viz., motor ability, attention, memory, imagination, analytic comparison, generalization, inference.

(II) The functioning of intellect :
  (1) In the command of language.
  (2) In the application of past experience to theoretical and practical problems.
  (3) In the use of standards of judgment and in the quality of judgments.

(III) The interest determinants of theoretical and practical ability.

The Binet-Simon tests in their original and adapted forms, and 'alpha" test of the American Army, the College entrance tests, the Hamburg Secondary School selection tests, in so far as they were not information tests, fell under groups I and II. The educational bearing of the psychological doctrines of instinct and interest was not adequately recognised in existing tests, but was of first-rate importance in interpreting results (Group III).

To a great extent interest and ability went together, and it was almost impossible to test one without the other. A child was born with certain range of instinctive ability varying in degree in different directions, such ability determining interest, but the child had also spontaneous activity and from this as well as from his instincts developed the interests which made up his individuality. Unless ability was tested along the lines of interest the child had not a fair chance of making a good score.

Special and Group Abilities.

The whole question of "special" and "group" abilities was still in the melting-pot. The validity of inferences based on correlation co-efficients as to group abilities or a general common factor was questionable. The conditions upon which excellence of performance in a particular set of tests depended were insufficiently known. Theories as to group abilities, levels, transfer of training, application of a common method, appeal to a common interest, needed further elucidation. In the present state of ignorance a varied range of tests was essential.

A definite programme of tests extending over a considerable period of time and carried out on an extensive scale was necessary for solution of this and other problems.

The late Mr. J. A. GREEN, Professor of Education at the University of Sheffield.

THE NATURE OF INTELLIGENCE.

Investigation had so far been based upon two assumptions :—

(*a*) That intelligence was something in the nature of mind energy (a central factor), which entered, in varying degrees, into all controlled behaviour, from simple repeated movement through sensory discrimination to the recognition of analogies and opposites, from the most elementary function (single factors isolated as far as possible) to the most complex function (group factors working in combination).

(*b*) That a hierarchical system of correlations was evidence of the existence of a central factor.

On the first point one might perhaps urge the biological point of view as more likely to help us in deciding upon the nature of intelligence. It seemed from that standpoint to be best defined as the integrative function of mind. In the last resort this integrative function found expression in adaptive behaviour. (He used the word integrative, rather than integrating, in order to avoid confusion with the mathematical use of the word. Here he meant the capacity for seeing relations, for systematisation if thought better, but not summation.) Failure to integrate one's world meant incoherence in one's behaviour within that world. If general ability meant intelligence as he had defined it, it was a single function operating "at call," but with varying efficiency. It ought to integrate, to unify experience, but it operated at various cognitive levels.

On the second point the question whether a hierarchy of correlations could be artificially produced where there was no central factor at work, and whether the mere discovery of a hierarchy was, therefore, conclusive evidence of the existence of a central factor, were under dispute and the witness was awaiting the outcome.

At this point one might perhaps usefully draw a distinction between intellect and intelligence. The biologist distinguished tacitly between them when he spoke of an order of intelligence in the animal world. A dog might be very intelligent, but never intellectual. A workman was often very intelligent though even under the most favourable environment he would not have been intellectual. Conversely, highly intellectual men were sometimes less intelligent than they were intellectual. Gaucherie and absent-mindedness, and even childishness, were features commonly expressed in the caricature of the professor. This might, of course, be expressed in another way. Mind never worked in a vacuum. It might create a world of its own and determine behaviour with reference to that

world. Within this intellectually created world, the professor was highly intelligent, but he had in the process of building his world lost touch with that of ordinary men, in which world therefore he often behaved foolishly.

If this view were true, it seemed to follow—

(*a*) That intelligence was not a summation of parts, or a multi-dimensional entity capable of resolution in certain directions.

(*b*) That intelligence was correlative to the universe in which it worked. As the universe developed, intelligence developed. They conditioned each other, and intelligence was revealed by its universe of operation. In other words one could not measure a potential intelligence, though one might, in some sense, gauge an active intelligence by a comparative estimate of its efficiency within its sphere of work. Intelligence did not exist *in vacuo*. Professor Green did not agree with the expert psychologists who said that the intelligence of the individual did not grow after about the age of 15 or 16. He considered that as a man's environment expanded so his intelligence grew. He would accept the statement that a man of 40 would do no better than a boy of 16 in certain psychological tests (*e.g.*, in the detection of absurdities), but he would not accept such tests as a measure of intelligence. The man of 40 required a different test of intelligence from the boy of 16. If we are to give both the same test, it must obviously be a test within the universe of the 16 year old. That meant a serious handicap for the older man. It was perhaps comparable to a test devised to compare the intelligence of a child and a dog. To make comparison possible the test must have reference to a world in which both child and dog could function. Might not the dog prove superior in such a case ?

(*c*) That tests of intelligence must take into chief consideration the world in which the persons concerned had been living, and the possibilities of their rising to the demands of a larger world.

Thus the comparison of one intelligence with another involved the consideration of at least four factors—the two universes, and the two efficiencies within those universes.

For example :—A in a universe *a*, had efficiency ·75.
B in a universe *b*, had efficiency ·5.

The universe *b* was, however, *x* times more complex than the universe *a*. Then, what was the relation of the intelligence of A to that of B ? Whatever answer one might give to this problem did

not meet the whole need, for A in the simple universe *a* might not be stretched to his full power. His efficiency in universe *c*, which was *y* times as complex as *a*, might actually rise to ·85—a situation not uncommon in comparing results in the universe of school with those in the universe of life. One wanted, in fact, something in the nature of a co-efficient of elasticity in making comparisons amongst intelligencies. This might be a constant in a given individual. If so, it was the most important of all.

THE NEED FOR TESTING FOR "SPECIAL" AND "GROUP" ABILITIES:

With regard to the need for testing "special" and "group" abilities, Professor Green said that if this meant the various types of universe in which different "intelligence" might operate most efficiently there was probably real need for enquiry in this direction. Psychological tests were probably used very slightly anywhere except in America.

There were some practical difficulties in assessing intelligence (arising from its mode of working).

(*a*) Intelligence was often curiously shy. Even in the same person its operative force could not be steadily guaranteed. This was shown by the relatively low reliability co-efficients often obtained in various kinds of problem work.

(*b*) Intelligence was apt to show itself in restricted fields, or to operate with varying efficiency in different fields, *e.g.*, a class of fifty boys (Standard VA) working arithmetical problems and Burt's reasoning tests gave a correlation as low as ·34.

There were many schoolboys (perhaps more school girls) who were prevented from matriculating by the requirements in Mathematics.

These differences were probably due to fundamental differences (*e.g.*, simple quantitative relations) in sensitivity. Restriction of this kind, however, meant an intelligence of restricted range, and a test in a wrong field would give a very unreliable result for the individual concerned.

(*c*) Even though intelligence was defined as the integrative function of mind the integrative process really included analysis. It could only put together what it has previously taken to pieces, but the taking to pieces (when it was conscious), must have been done upon a plan which itself was the outcome of an integral view, and when it was not conscious, as in childhood, the analysis followed more or less closely, a plan determined by our

racial history. The difference between a penetrating and a superficial intelligence was due to differences in analytical power. So we were brought back again to the capacity to see things as wholes as being a fundamental characteristic of intelligence.

Mr. E. O. Lewis, D.Sc., M.R.C.S.

Factors of General Ability.

Dr. Lewis said that all tests of general ability were really tests of special abilities. Although there is a " common factor " involved in general ability, no tests yet devised measure this factor directly. The most important factors involved in general ability are voluntary attention, reasoning, the ability to choose the essentials in any situation, and the power to apply previous knowledge in solving new problems.

One method of approaching this problem was to find what mental tests differentiated most definitely the normal from the subnormal person. Records with the following Binet tests showed the greatest disparity between these two classes of persons :—

Comprehension of difficult questions.
Reconstruction of dissected sentences.
Detection of absurdities.
Definition in terms superior to use.
Definition of abstract terms.
Tests of immediate memory.
Counting backwards.

These results supported Binet's view of the importance of linguistic ability in the assessment of general intelligence. The failure of subnormal persons in the tests for immediate memory witness believed to be due chiefly to defective concentration of attention ; and their failure in counting backwards to inability to manipulate mental associations previously acquired. These opinions were based on investigations, lasting over many months, that he had carried out amongst mentally defective children. The poor memories of defective children were due not so much to subnormal physiological retentiveness as to their inability to adopt systematic and rational method of learning. There were cases, however, where defective children possessed extraordinary memories, but such children often failed hopelessly with such practical tests as Healy's Construction Board.

Special and Group Abilities.

The present system of educating children in large classes made it difficult to recognise special abilities, and still more difficult to foster them. Specially devised psychological tests could give the teacher valuable information in this respect.

In the past, the failure to analyse carefully the child's mental processes had led us to attribute to general intelligence much that was really a manifestation of special abilities.

In the study of mentally defective children *specific disabilities* seemed to be of greater significance than the general mental retardation.

Mr. B. Muscio, formerly Investigator to the Industrial Fatigue Research Board (now Challis Professor of Philosophy at the University of Sydney).

"General" Ability.

Mr. Muscio suggested that educable capacity might not be a unitary thing, but that a person might be educable in one direction and not in another. Sufficient investigation had been done to justify persistent research into this suggestion. Probably, however, educable capacity of any sort demanded the presence of the following three factors, involved in "general" ability.

(*a*) The capacity to perceive or apprehend relations.

(*b*) The capacity to adjust thinking and action to new conditions; and

(*c*) The capacity to combine particular experiences into significant wholes.

The two latter factors might be dependent upon the first, which would then be the fundamental factor concerned. But the capacity to perceive relations was possibly specialised. There was the person who could best perceive relations in concrete objects, such as machines, and the person who could best perceive relations among ideas expressed in words. Possibly those cases indicated different levels of "general" ability rather than different *kinds* of it.

Mr. T. Percy Nunn, D.Sc., Principal of the London Day Training College and Professor of Education in the University of London.

Theory of Tests.

Dr. Nunn stated that in testing a person the aim was at bottom, always the same; namely, to infer from the way in which he dealt with a given situation the way in which he was likely to deal with others. The inference was safer the more closely the subsequent situations resembled the first, but it was never free from an element of risk. A boy might jump a certain height to-day, but it was not certain that he would be able to reproduce the feat to-morrow. In general, however, the situations contemplated in the inference were of necessity different from the test-situation, and might differ

from it very widely—as when from the quality of a young man's Greek verses his probable capacity as an administrator was inferred. In such cases the inference was valid only if it were known beforehand that the capacities brought into play in the test-situation and in the subsequent situations were, if not identical, at least definitely correlated. In other words, the procedure must be based upon a well-founded theory of abilities.

At present there was no such theory universally accepted by psychologists. It might be taken as agreed that situations could be classified into groups, such that inference was always possible from observed ability to deal with one situation to probable ability to deal with another of the same group. For instance, a boy who was good at tennis was very likely to be good at racquets, fives and badminton; and one who learnt Latin easily would probably learn certain other languages easily. The question in dispute was whether ability in one group was a valid basis for inferring ability in a group which appeared to have nothing in common with the former. To that question Professor Thorndike has replied "No," Professor Spearman, "Yes."

It was somewhat widely believed that Professor Godfrey Thomson who had dealt with the subject very acutely, had recently proved Dr. Spearman's answer to be untenable. This was a mistake. Spearman had expanded his "Yes" into the statement that a single "central intellective factor" entered into the constitution of all varieties of intellectual performance; it was against this statement, and, in particular, against the chief mathematical argument used to support it, that Thomson had directed his attacks. It must not, however, be assumed that these attacks had been successful. From personal discussion, Professor Nunn had gathered that Professor Thomson did not dispute the position in the form in which witness had expressed it; namely, that a person's powers to deal with situations of widely differing character were, in so far as they involved intellect, connected systematically in such a way that it was always possible to make an inference from his ability in dealing with one type to his probable ability in dealing with another type. And that Professor Nunn took to be the essence of Professor Spearman's contention. He himself agreed with this contention.

The controversy had far more than mere academic interest; for unless Spearman's view, as Professor Nunn had put it, was sound, inferences based upon the results of tests—whether the new psychological tests or the older competitive examinations—must often be sadly precarious.

He agreed that there were sufficient instances of people being abnormally gifted in one activity and far below normal in all others to make the subject an important one for discussion. At present

there was no explanation; a partial explanation might be that the one subject had sapped all the individual's interest, leaving none for other subjects. There was also evidence of people whose interest took one direction at first and another direction later. With regard to great musicians witness considered that ability in music must also presuppose general ability.

He would define precocity in children as the premature emergence of certain faculties as compared with the average time of emergence in normal children. He was not prepared to say that a child could be precocious in one form of ability and of less than average ability in other forms; it was probably more true to say that precocity meant more general capacity at a certain age.

Mr. C. A. Richardson, H.M. Inspector of Schools.

Note.—It is to be understood that the opinions expressed by Mr. Richardson are his private views, and must not be taken to represent the opinion or the policy of the Board of Education.

The Factors Involved in "General" Ability which must be included in Testing for Educable Capacity.

Mr. Richardson said that while it should not be forgotten that in all mental processes the mind worked, as a whole, and therefore any isolation of special factors was largely arbitrary and abstract, the various *aspects* of intelligence might be conveniently classified so far as testing for educable capacity was concerned, as follows :—

(1) Readiness and ability in applying knowledge.
(2) Discrimination of essentials.
(3) Associative processes : (*a*) richness and maturity ; (*b*) logical integrity.
(4) Power to control and concentrate attention.
(5) Power of comprehension.
(6) Ability to hold in mind the conditions of a problem.
(7) Practical judgment and ingenuity.
(8) Steadiness of purpose.
(9) Power of forming abstract ideas.
(10) Power of generalisation.
(11) Critical ability, including self-criticism.
(12) Ability to manipulate imagery.

The items on the above list were, of course, far from being mutually exclusive ; there was considerable overlapping. Attention, for example, entered into all "intelligent" mental processes, and association into nearly all. But it was possible to say of any well-devised intelligence test, that it tested certain aspects of intelligence more particularly than others, and a set of tests should certainly be

constructed in such a way as to afford evidence of the child's capacity in all the above directions, if valid conclusions were to be drawn as to his educable capacity.

The Need for Testing "Special" and "Group" Abilities in Determining Educable Capacity.

Mr. Richardson thought it unlikely that special abilities existed in the same sense as general ability. It was becoming sufficiently evident that in all forms of "intelligent" mental activity there was present a certain factor to which the name "general ability" might fairly be given. It then seemed probable that special ability or (better) "aptitude," was a combination of general ability and special *interest*, and, in some cases, special temperamental and physical characteristics.

Hence in determining educable capacity the prime necessity was to determine the level of general ability. In children of superior intelligence this would usually be sufficient, for as a result of the tests they would be known to be easily educable, and any special aptitude would, without doubt, show itself in the natural course of their development. But in the case of children of a low degree of intelligence, it would be extremely useful to discover their special aptitude (if any), for it would then be possible to make the most of their limited educable capacity.

Briefly then, supposing all children to be tested for general ability, it was desirable in the case of those testing below a certain level, to make further suitable tests in order to discover their special aptitudes.

Mr. Godfrey H. Thomson, D.Sc., Ph.D., Professor of Education at Armstrong College, Newcastle-on-Tyne.

General Ability.

He was not personally a believer in the existence of a "faculty" called general ability. He considered that the statistical work of those who affirmed this theory was of doubtful validity. He had, of course, no objection to the popular use of the term to indicate the average ability which a man showed.

His own opinion was that ability in man was a very much more complex affair than this, and he was convinced that differences of type occurred which were not explicable by a general factor plus a unique or specific factor. It was not necessary to postulate the existence of a general factor, but educational units should be considered. The fact that a man was better in any particular activity was due to the pruning away of unnecessary mental factors.

Mr. Frank Watts, Assistant Inspector of Schools, formerly Lecturer in Psychology in the University of Manchester and in the Department of Industrial Administration of the College of Technology, Manchester.

Note.—It is to be understood that the opinions expressed by Mr. Watts are his private views, and must not be taken to represent the opinion or the policy of the Board of Education.

Factors Involved in "General" Ability.

The ability that became and remained "general" was probably rare. "General ability," as the psychologist knew it, was the form of intelligence which was of the nature of intellect as distinct from instinct or intuition—it involved the capacity to deal with "general" ideas. Tests *like* the U.S.A. Army *Alpha* tests would discover general intellectual ability. The *young* child had general intelligence rather than "general ability"—if one might be allowed to differentiate; it was general ability undeveloped, still at the concrete level, which he displayed. The only method of testing general intelligence at this stage would seem to be to take samples of ability in a very large number of diverse performances. This was impossible practically. A sampling would usually introduce the element of unfairness.

"Special" and "Group" Abilities.

The "special" abilities which needed investigation were :—

I.

(1) *Musical Ability.*—Witness recommended the use of the Seashore Gramophone Test Records for this purpose. (Ability to discriminate differences of pitch and intensity, sense of consonance and tonal memory were tested.)

(2) *Arithmetical Capacity.*—The ordinary examinations usually did this, but they were occasionally apt not to produce thoroughly reliable results.

(3) *Language Capacity.*—This was usually tested *via* Composition. Added information could be secured from children's *gradings* of a number of Compositions of obviously varying merit.

(4) *Drawing Capacity and Mechanical Dexterity.*—Witness knew of no reliable tests.

II.

Arising from the above and energised by particular "instincts" and "interests" are many of the aptitudes making for success in industry, trade and commerce. At present these are untestable.

Mr. W. H. WINCH, District Inspector of Schools under the London County Council.

GENERAL ABILITY.

Mr. Winch said that the meaning and existence of general ability was much disputed among psychologists. The older conception of faculties was largely given up. Personally he would include a test in Rote Memory, Substance Memory, Imagination and Reasoning as necessary to estimate so-called general ability, with possibly a Perception test. There was a difference between " rote memory " and " substance memory " : the one entailed the memorising and reproduction of words and the other the reproduction of ideas underlying the passage to be memorised. Of the two " substance memory " was usually the better done, as it was the easier. The ability for rote memory and for substance memory steadily rose up to the end of the elementary school age.

Whether abilities did fall into " groups " was again a matter of dispute among psychologists. Using the term " general ability " loosely as a sum or average of several factors, witness thought that tests should be used in all cases, where possible, as above indicated. Special tests would be required for draughtsmanship and constructive ability of a mechanical kind in addition to those of linguistic constructiveness involved in the general ability tests.

INDEX.

PAGE

ABSURDITIES.—Specimens of 17, 216
ACTON AND CHISWICK JUNIOR TECHNICAL SCHOOL.—Use of Individual Tests at 159
AESTHESIOMETER.—Use of, in Testing for Touch Discrimination 7, 101
ALGEBRA.—Tests in 46
ALTONA.—*See* Germany.
AMERICA, UNITED STATES OF.—Army—Use of Tests in .. 36, 50, 179
Specimens of Tests used in 210
Binet-Simon Scale, Revisions of 30
Boston Vocation Bureau—Work of 48
Schools, Note on Grades in 190
Use of Tests in, (General) 26, 37, 178
Performance Tests 42
Standardised Scholastic Tests 44, 45
Vocational Tests 48, 182
AMSTERDAM.—*See* HOLLAND.
ANALOGIES.—Specimens of 16, 208
Use of 88
ANATOMICAL STIGMATA.—*See* PHYSICAL TESTS.
ANDERSON, Miss G. M.—Use of Individual and Group Tests by .. 157
APPLICATION OF TESTS.—Training of Persons who apply Tests 105, 139
APPRENTICE SCHOOLS.—Vocational Selection for 52
AQUINAS, ST. THOMAS.—Definition of "Intelligence" 68
ARISTOTLE.—On Distinction between "General" and "Special" Intelligence 15
ARITHMETIC.—Tests in 45, 46, 88, 220
ASSOCIATION TESTS.—Use of 88
ATTENTION.—Use of Tests of 131, 144
AUSTRIA.—Use of Tests in 164
AYRES, Dr. L. P.—Scale of Tests in Handwriting by 45
Spelling by 45

BALLARD, Dr. P. B.—Absurdity Tests by 17
Arithmetic Tests by 46
Specimens of Chelsea Tests by 215
Use of Tests by 158
See also CHELSEA TESTS.
BALTIMORE.—*See* AMERICA.
BARCELONA.—*See* SPAIN.
"BATTERIES" 207
BELGIUM.—Use of Tests in 51, 165
BERLIN.—*See* GERMANY.
BERNE.—*See* SWITZERLAND.
BERTILLON, Monsieur.—Determination of "Physical" Constants by 23
BIBLIOGRAPHY.—List of Recent Publications in Psychology and Psychological Tests 191

PAGE

BINET, Dr. A.—Correlational Methods by 13
" Intelligence "—Conception of 69
Tests by—for Measuring Re-action Time 10
Standardised Scholastic Tests.. 45
of Temperament and Character 55, 56, 93
of Touch Discrimination 7
(*See also* BINET-SIMON SCALE.)
BINET-SIMON SCALE—Examples of, ... 200
History of 22
Limitations of 40, 76, 142
List of Works on 193
Measurement of Intelligence by 27
Merits and Disadvantages of 78
Revisions of—
by Goddard 30
the London Revision 32
the Stanford Revision 31, 77, 84
the Terman Extension 84
the Treves Saffiotti Method 33
the Vineland Revision 30
the Yerkes Point Scale 30, 83
Use of, for Blind and Deaf Children 78
Junior Scholarships 78
Mentally Deficient Children 23, 78, 82
Super-normal Children 78
BINGHAM, Dr. W. V.—Testing in American Army by 51
BIRMINGHAM.—Use of Tests in the Administration of Justice .. 54
BLACKPOOL.—Use of Group Tests for Junior Scholarship Purposes at 38, 149
BLIND CHILDREN.—Use of Tests for 78
BORDEAUX.—*See* FRANCE.
BOSTON.—*See* AMERICA.
BOURDON, Monsieur.—Use of Tests of Vision by 8
BRADFORD.—Use of Group Tests for Junior Scholarship Purposes at 37, 152
BRESLAU.—*See* GERMANY.
BRISTOL UNIVERSITY.—Use of Tests at the Department of Education of 39
BRNO (BRÜNN).—*See* CZECHOSLOVAKIA.
BROWN, Dr. W.—Views on " General " Ability 225
Physical Tests.. 101
BRUGMANS, Dr.—Use of Tests by 173
BRUSSELS.—*See* BELGIUM.
BUCKINGHAM, Dr. B. R.—Tests in Spelling by 45
BURT, Dr. CYRIL.—Bibliography by 191
Correlation, Note on 185
Development of Psychological Tests—Historical Sketch of .. 1
Examples of Tests Selected by 199
London Revision of Binet-Simon Scale by 32, 203
" Intelligence," Views on 15, 70, 226
Investigations at Oxford and Liverpool 15
Norms of Performance, Views on 96, 183
Physical Tests, Views on 102
Standardised Scholastic Tests, Views on 96, 183

CANAL BOAT CHILDREN.—Mental Testing of, by Mr. Hugh Gordon.. 40
CANCELLATION TESTS.—Description of.. 8

PAGE

CATTELL, Prof.—Researches of 5
Tests of Temperament and Character by 55
CENTRAL COMMITTEE.—*See* CLEARING HOUSE.
CENTRAL SCHOOLS.—Use of Tests for Admission to 78, 111, 141
Use of Tests in 127, 142
CHARACTER.—Importance of 73
Tests of 53, 93, 132, 144
List of Works on 197
CHARLOTTENBURG.—*See* GERMANY.
CHELSEA TESTS.—Description of 17, 46
Specimens of 215
Use of 158
See also BALLARD, Dr. P. B.
CHELTENHAM, PATE'S GRAMMAR SCHOOL.—Use of Group Tests at 38, 160
CHEMISTRY.—Tests in 46
CHICAGO.—*See* AMERICA ..
CHINA.—Use of Tests in 165
CHRISTIANIA.—*See* NORWAY.
CIPHER TEST.—Specimens of 216
CIVIL SERVICE EXAMINATIONS.—Use of Group Tests in 39
CLEARING HOUSE.—Necessity for 106, 135, 144
CLEVELAND (OHIO).—*See* AMERICA.
COACHING.—Advantages and Disadvantages of Special Preparation for Tests 103
CLOCK TESTS.—Inclusion of, in the Vineland Revision 30
CODE TESTS.—Inclusion of, in the Vineland Revision 30
COLLECTIVE TESTS.—*See* GROUP TESTS.
COMPLETION TESTS.—Invention of 25
Specimens of 209
Use of 88
COMPOSITION.—Tests in 46
CONSTRUCTION OF TESTS.—To be Entrusted to Trained Psychologist 105, 136
COPENHAGEN.—*See* DENMARK.
CORBERI, Professor.—Use of Tests by 174
CORRELATION.—Degree of, between Tests and other Examinations 38, 39
Explanation of 13, 185
COURTIS, Professor S. A.—Arithmetic Tests by 46
CUBE CONSTRUCTION TEST.—Specimen of 206
CUBE IMITATION TEST.—Specimen of 206
CZECHOSLOVAKIA.—Use of Tests in 166

DEAF CHILDREN.—Use of Tests for 78
DEFECTIVE CHILDREN.—*See* MENTAL DEFICIENCY.
DELINQUENT CHILDREN.—Testing of 53, 54, 199
DENMARK.—Use of Tests in 166
DENVER.—*See* AMERICA.
DE SANCTIS, Professor TAUTE.—Construction of Tests by .. 34
Use of Tests by 86, 173, 174
DIAGONAL TEST.—Description of 206
DIFFERENTIAL PSYCHOLOGY.—Origin of 1
DOBSON, Mr. R. R.—Use of Tests at Cheltenham Grammar School by 38, 160
DRAWING.—Tests in 46
DRESSMAKING.—Vocational Tests in 223

PAGE
DREVER, Dr. JAMES.—Views on "General" Ability 227
Views on Standardised Scholastic Tests 97
DRUMMOND, Dr. W. B.—Use of de Sanctis Tests by 35

EBBINGHAUS, Dr. H.—Use of the Kombinations-Methode (Completion-Method) Tests 25, 88
EDGELL, Dr. BEATRICE.—Views on "General" Ability and "Special" or "Group" Abilities 228
ELEMENTARY SCHOOLS.—Use of Group Tests in 19, 125, 142
Use of Individual Tests in.. 125
Use of Standardised Scholastic Tests in 130
Use of Vocational Tests in 130
EMOTION.—Connection between Emotion and General Educable Capacity 95
Testing of 54, 55
ENVIRONMENT.—Effect of, on "Intelligence" 75, 138
EXAMINATIONS.—Aims of 109
Classification of Various Kinds of 109
Connection of Tests with Ordinary Examinations .. 108, 140
Defects of 63, 65, 115
First and Second School Examinations 111, 138
Free Place Examinations 108, 113, 118, 141
Objections to any Scheme of Examination which Determines Educational Future of any Child at 11 years of age .. 108, 138
Oral Examinations 120, 132, 141, 160
Stereotyped 44

FINLAND.—Use of Tests in 167
FONTÈGNE, Prof. J.—Use of Vocational Tests by 168
FRANCE.—Use of Tests in 25, 45, 167
Use of Standardised Scholastic Tests in 45
Use of Vocational Tests in 167
FREE PLACE EXAMINATIONS —*See* SCHOLARSHIPS

GALL.—System of Phrenology by 2
GALTON, Sir FRANCIS.—Researches into Mental Testing 5, 7, 20, 55, 56
GALVANOMETER.—Use of.. 57
GEMELLI, Professor.—Use of Tests by 174
GENERAL ABILITY.—Burt's Verification of the Theory of 15
Spearman's Theory of 14, 225
Various Hypotheses Regarding 68, 225
See also INTELLIGENCE.
GENERAL FACTOR.—*See* INTELLECT.
GENEVA.—*See* SWITZERLAND.
GEOGRAPHY.—Tests in 46
GERMANY.—Use of Tests in 25, 51, 168
GILBERT, Professor.—Use of Tests by 7, 8, 10
GIPSY CHILDREN.—Testing of, by Mr. Hugh Gordon 40
GODDARD, Dr. H. H.—Revision of Binet-Simon Scale by 30
GORDON, Mr. HUGH.—Mental Testing of Gipsy and Canal Boat Children 40
GOTHENBURG.—*See* SWEDEN.
GREEN, the late Professor J. A.—Investigation into Standardised Scholastic Tests by 46
Views on the Nature of "Intelligence".. 229
GRONINGEN.—*See* HOLLAND.

PAGE

GROUP TESTS OF INTELLIGENCE.—Comparative Merits and Disadvantages of, as Compared with Individual Tests 90
Specimens of 207
Training of those Engaged in the Administration of 105
Use of (General) 19, 35, 37, 87, 138, 141, 142
at Rugby School 35, 162
for Adults 39
in American Army 36, 50, 179
at Blackpool 38, 149
at Bradford 37, 152
in Civil Service Examinations 39
in Free Place and Junior Scholarship Examinations 37, 38, 114, 118, 119, 126, 141, 149
in Leicestershire 152
in London 157
in Northumberland 38, 153
in Yorkshire, West Riding 153

HALLE.—*See* GERMANY.

HAMBURG.—*See* GERMANY.

HANDWRITING.—Ayres' Scale for 45
Thorndike's Scale for 45

HARVARD UNIVERSITY.—*See* AMERICA.

HEALY PUZZLE TEST.—Description of 206

HEARING, TESTS OF.—*See* SENSORY TESTS.

HELSINGFORS.—*See* FINLAND.

HISTORY.—Tests in 46

HOLLAND.—Use of Tests in 173
Use of Vocational Tests in 51, 173

IKIN, Dr. A. E.—Use of Group Tests and Standardised Scholastic Tests by 38, 149

INDIVIDUAL PSYCHOLOGY.—Origin of 1

INDIVIDUAL TESTS.—The Binet-Simon Scale 22, 143
Comparative Merits and Disadvantages of as Compared with Group Tests 90
Place of, in Interviews 92
Specimens of 200
Training of those Engaged in their Administration .. 105, 139
Use of in Free Place Examinations 152, 157
Use of in Schools 159, 189

INDUSTRIAL FATIGUE RESEARCH BOARD.—Experiments in Performance Tests by 43
Experiments in Vocational Tests by 51

INTELLECT.—Existence of Intellectual Factors 19, 225

INTELLIGENCE.—Meaning of, as Used by Psychologists 67, 137, 225
See also GENERAL ABILITY.

INTERPRETATION OF RESULTS OF TESTS.—Training of those Interpreting Results 105, 134, 136

INTERVIEWS.—
Place of Individual Tests in 92
Use of Manual Tests in 93
List of Works on 198
See also ORAL EXAMINATIONS.

ISAACS, Mrs. S. S.—Views on the Connection between Emotion and General Educable Capacity 95

ITALY.—Use of the Doctrine of Anatomical Stigmata by the Italian School of Criminology 2
Use of Tests in 173

PAGE

JAPAN.—Use of Tests in 174
JONES, Dr. CARADOC.—Physical Tests by 102
JUNG, Dr. C. G.—Psycho-Analysis by 56
JUNIOR TECHNICAL SCHOOLS.—Use of Tests in .. 127, 142, 159

KALAMAZOO.—*See* AMERICA.
KELLEY, Mr. F. J.—Tests in Reading by 46
KRAEPELIN, Dr.—Researches of 13

LATIN.—Tests in 46
LAVATER.—Ideas of, on Physiogonomy 2
LEICESTERSHIRE.—Use of Tests in Examinations for Free Places in .. 152
LEWIS, Dr. E. O.—Views of, on "General" Ability and "Special" and "Group" Abilities 232
LINGUISTIC TESTS.—Limitations of 40
LOMBROSO, CESARE.—Work of, on Anatomical Stigmata 3
LONDON.—Use of Group Tests at Mile End Central School 159
Use of Individual and Group Tests at the Ben Jonson Council School, Stepney 157
Use of Performance Tests in 42
Use of Standardised Scholastic Tests in 47
Use of Tests at Bedford College 39
London Day Training College 39
University College 39
LONDON REVISION OF BINET-SIMON SCALE.—Description of .. 32
Specimens of 203
LUXEMBURG.—Use of Tests in 174

McDOUGALL'S PLUNGER.—Use of 101
MANCHESTER.—Use of Physical Tests at 102
MANUAL TESTS.—Use of in Interviews 93
Use of in Vocational Testing 98, 130
List of Works on 196
MARKING RESULTS OF TESTS.—Training of Persons to Mark Results .. 105
MEMORY.—Specimens of Tests of 218
Use of Tests of 131, 144
MENTAL AGE.—Constancy of 65, 85
Construction of a Scale of, by Binet 27
MENTAL DEFICIENCY.—Diagnosis of, by Means of Binet-Simon Scale 23, 82
Use of Tests for Determining ..41, 126, 134, 139, 140, 178, 198
MENTAL IMAGERY.—Galton's Researches into 20
Use of Tests of 131
MILAN.—*See* ITALY.
MIXED RELATIONS TEST.—*See* ANALOGIES.
MODENA.—*See* ITALY.
MORALITY.—Tests of 58
MOTOR TESTS.—Correlation between Intelligence and Speed of Re-action Time 10, 16
Measurement of Re-action Time 10
Use of 9, 18
MÜLLER, Mr. E. K.—Discovery of Psycho-Galvanic Re-action by .. 56
MUMFORD, Dr. A.—Physical Tests by 102
MUNROE, Mr. W. S.—Tests of Reading by 46
MÜNSTERBERG, Professor HUGO.—Experiments in Vocational Selection by 49
MUSCIO, Professor B.—Views on "General" Ability 233
MUSCLE SENSE.—*See* SENSORY TESTS.

PAGE

NANTES.—*See* FRANCE.
NATIONAL INSTITUTE OF INDUSTRIAL PSYCHOLOGY.—Aims of .. 51
Vocational Tests by 221
Work of 51
NATIONAL INTELLIGENCE SCALE.—Use of in America 36
Use of at Blackpool 38
Use of in Northumberland 154
NEUROTIC CHILDREN.—Testing of 53
NEW HAVEN.—*See* AMERICA.
NEW ROCHELLE.—*See* AMERICA.
NORMS OF PERFORMANCE.—Necessity of 96
Note on, by Dr. Cyril Burt 183
NORSWORTHY, Dr. NAOMI.—Tests of Temperament and Character by 55
NORTHUMBERLAND.—Use of Tests for Junior Scholarships in.. 38, 153
NORTHUMBERLAND TESTS.—Use of, for Junior Scholarships .. 38, 153
Examples of 214
See also THOMSON, Professor GODFREY H.
NORWAY.—Use of Tests in 174
NUNN, Professor PERCY.—Views of, on Norms of Performance .. 97
Views of, on Theory of Tests 233

OAKLAND (CALIFORNIA).—*See* AMERICA.
OEHRN Dr.—Use of Tests of Visual Discrimination by 8
OPPOSITES.—Examples of 207
ORAL EXAMINATIONS.—Further Research Desirable 141
Importance of 120, 132, 141
List of Works on 198
Use of Intelligence Tests in 120, 160
Use of, in Free Place Examinations 121
See also INTERVIEWS.
ORIENTATION TESTS.—Specimens of 217
OSWEGO (NEW YORK).—*See* AMERICA.
OTIS GROUP TESTS.—Use of in America 36
Use of in Blackpool.. 38

PARIS.—*See* FRANCE.
PARSONS, Mr. FRANK.—Experiments in Vocational Tests by .. 48
PASSIAC (NEW JERSEY).—*See* AMERICA.
PATERSON AND PINTNER.—Collection of Performance Tests by 41, 206
PEARSON, Professor KARL.—Researches into the Value of Cranial Stigmata 3
PERCEPTION.—Use of Tests of 131, 144
PERFORMANCE TESTS.—Specimens of 205
Use of 40, 41, 42, 66, 159
List of Works on 196
PHILADELPHIA.—*See* AMERICA.
PHRENOLOGY.—*See* PHYSICAL TESTS.
PHYSICAL TESTS.—Anatomical Stigmata 2
Correlation of with Mental Tests 101
Definition of 100
Gall's System of Phrenology 2
Lavater's Ideas on Physiognomy 2
Origin of 2
Use of 100, 131, 144
Value of 101
PHYSICS.—Tests in 46
PHYSIOGNOMY.—*See* PHYSICAL TESTS.

PAGE
PINTNER AND PATERSON.—Collection of Performance Tests by 41, 206
PITCH DISCRIMINATION.—Use of Tests of 9
PITTSBURGH.—*See* AMERICA.
PIZZOLI, Professor UGO.—Use of Tests by 173
PORTEUS MAZE TESTS.—Use of 41
PRACTICAL TESTS.—*See* MANUAL TESTS *and* PERFORMANCE TESTS.
PRAGUE.—*See* CZECHOSLAVAKIA.
PSYCHO-ANALYSIS.—Use of 56
PSYCHO-GALVANIC REACTION.—Discovery of, by Mr. E. K. Müller .. 56
PSYCHOLOGY.—List of Works on (General) 191
List of Works on (Educational) 192

RANGE OF INFORMATION TESTS.—Use of 88
RE-ACTION TIME.—*See* MOTOR TESTS.
READING.—Specimens of Tests of 219
Use of Tests of 45, 46, 89
REASONING TESTS.—Researches of Mr. W. H. Winch 18
Specimens of 17, 204
See also ABSURDITIES, ANALOGIES, SYLLOGISMS.
RETARDED CHILDREN.—Use of Tests for 126
RICE, Dr. J. M.—Investigation upon Spelling 44
RICHARDSON, Mr. C. A.—Views on "General" Ability and "Special" or "Group" Abilities 235
Use of Tests by 154
ROCHESTER (NEW YORK).—*See* AMERICA.
ROME.—*See* ITALY.
RUGBY SCHOOL.—Use of Tests at 38, 162

SAFFIOTTI, Professor U.—Modification of Binet-Simon Scale by .. 33
SCHOLARSHIPS.—Improvement in the Technique of Free Place Examinations 113
Objections to the Limitation of the Number of Free Places .. 108
Shortcomings of the Present Examinations for Free Places .. 115
Use of Tests for Free Place and Junior Scholarship Examinations 37, 38, 78, 114, 118, 119, 126 141, 149
Use of Oral Examinations for 121
SCHOLASTIC TESTS.—*See* STANDARDISED SCHOLASTIC TESTS.
SCHOOL MEDICAL OFFICERS.—Use of Tests by .. 133, 135, 139, 140
SCHUYTEN, Dr.—Use of Tests of Touch Discrimination 7
SCOTT, Colonel W. D.—Testing in American Army by 51
SEASHORE.—Views on Correlation between Pitch and Discrimination and Intelligence 9
SECONDARY SCHOOLS.—Increase in Accommodation necessary .. 109
Use of Group Tests for Entrance to 38, 111, 141
Group Tests for Junior Scholarships to 37, 38, 111, 141, 149
Standardised Scholastic Tests in 130
Tests in (General) 127, 142, 160
Vocational Tests in 130
SEGUIN, Monsieur EDOUARD.—Form-board Devised by 41
SENSORY TESTS.—Correlation between Mental Ability and Sensory Discrimination 7, 8, 9, 16
Use of 6, 7, 8, 18, 101
SHANGHAI.—*See* CHINA.
SHORTHAND.—Vocational Tests in 221
SILESIA.—*See* GERMANY.
SIMON, Dr. T. H.—*See* BINET, Dr. A., *and also* BINET-SIMON SCALE.

PAGE

SIMPLEX TESTS.—Use of, for Junior Scholarships Purposes at Blackpool 38
Use of, for Junior Scholarships Purposes in Northumberland .. 154
SPAIN.—Use of Tests in 51, 175
SPEARMAN, Professor C. E.—Conception of "Intelligence" by 70, 234
Distinction of Special Capacities by 19
Investigation into Correlational Methods 13, 14
Theory of General Ability 14, 225
Use of Auditory Tests by 9
Simple Sensory Tests by 7
Visual Tests by 8
SPECIAL FACTORS.—*See* INTELLECT.
SPECIFIC ABILITIES.—Distinction of 19
Specimens of Tests of 217
SPELLING.—Graded Exercises in 45
Investigations upon 44
Specimens of Tests in 220
SPIELMAN, Miss WINIFRED.—Tests for Dressmakers' Apprentices by 223
STANDARDISATION OF TESTS.—By Binet 28
Note on, by Dr. Cyril Burt 183
To be Entrusted to Experts 136
Views of Witnesses on 96
STANDARDISED SCHOLASTIC TESTS.—Connection of with Ordinary Examinations 141
History of 43
List of Works on 194
Specimens of 219
Use of, by Teachers 106, 133, 140
Use of in America 44, 45, 129
Elementary Schools 130, 152
England 46, 129, 140, 151
France 45
Secondary Schools 130
Views of Witnesses on 96
STANFORD REVISION OF THE BINET-SIMON SCALE.—Decription of .. 31
Merits and Defects of 84
STATISTICAL METHOD.—Application of, to Education 71, 83, 90, 114, 119, 125, 136
(*See* also NORMS OF PERFORMANCE *and* STANDARDISATION OF TESTS.)
STERN, Dr. WILLIAM.—Conception of "Intelligence" by 70
STRASBOURG.—*See* FRANCE.
STUTTGART.—*See* GERMANY.
SWEDEN.—Use of Tests in 176
SWITZERLAND.—Vocational Testing in 51
SYLLOGISMS.—Specimen of 17

TAGG, Mr. MAX.—Use of Individual Tests by 159
TAPPING TESTS.—Use of 10, 101
TAYLOR, Mr. F. W.—Experiments in Vocational Selection by .. 49
TEACHERS.—Use of Tests by 105, 106, 133, 140
Training of Teachers to administer Tests 105, 133, 138, 139
TEMPERAMENT.—Tests of 53, 93, 132, 144
List of Works on 197
TERMAN, Professor L. M.—Revision of Binet-Simon Scale by 31, 84
Use of tests by 154
Use of Terman Group Tests at Blackpool 38
in America 36
in London 158

PAGE

THOMSON, Professor GODFREY H.—Conception of "Intelligence" by 70, 226, 234, 236
Use of Tests by 153
See also NORTHUMBERLAND TESTS.
THORNDIKE, Professor E. L.—Researches of 5, 11, 45, 46, 50
Theory of, Regarding "Intelligence" 69, 234
TIBBEY, Mr. J. G.—Use of Group Tests by 159
TOKIO.—*See* JAPAN.
TOMLINSON, Mr. T. P.—Use of Group Tests by 155
TOUCH, TESTS OF.—*See* SENSORY TESTS.
TRADE SCHOOLS.—Vocational Selection for 52
TREVES, Professor Z.—Modification of Binet-Simon Scale by .. 33
TREVES-SAFFIOTTI METHOD.—Description of 33
TRIANGLE TEST.—Specimen of 206
TWENTYMAN, Mr. A. E.—Note on Grades in American Schools by .. 190
TYPISTS.—Vocational Tests for Shorthand-Typists 221

UNITED STATES OF AMERICA.—*See* AMERICA.

VANEY, Monsieur V.—Standardised Scholastic Tests by 45
VAUGHAN, Mr. W. W.—Use of Group Tests at Rugby School by 38, 162
VIENNA.—*See* AUSTRIA.
VINELAND REVISION OF BINET-SIMON SCALE.—*See* BINET-SIMON SCALE.
VISION, TESTS OF.—*See* SENSORY TESTS.
VOCABULARY TESTS.—Specimens of 219
Use of 88
VOCATIONAL TESTS.—Administration of, and Interpretation of, Results of 106
Aims of 47
History of 47
Necessity of Testing for "General" Ability in 99
Specimens of 221
Use of 50, 98, 100, 130, 143, 159, 164
Performance Tests in 42, 159
Physical Tests in 101
Views of Witnesses on 99
List of Works on 197

WAGNER, Dr.—Use of Tests of Touch Discrimination by 7
WASHINGTON.—*See* AMERICA.
WATTS, Mr. FRANK.—Views on "General" Ability, and "Special" and "Group" Abilities 237
Test of Character and Morality, Experiments in 58
WHIPPLE, Professor J. M.—Use of Tests of Pitch Discrimination by 9
WHITLEY, Dr. MARY.—Statement as to the Use of Tests in America 179
WINCH, Mr. W. H.—Researches of 18, 46
Views on "General" Ability 238
WISSLER, Professor.—Researches of 12
WOODWORTH, Professor R. S.—Researches of 12

YERKES POINT SCALE.—Description of 30
Merits and Defects of 83
YORKSHIRE.—Use of Group Tests in Elementary Schools in the West Riding 155

ZÜRICH.—*See* SWITZERLAND.

Printed by H.M.S.O. Press, Harrow.

UNIVERSITY OF CALIFORNIA
MEDICAL CENTER LIBRARY

THIS BOOK IS DUE ON THE LAST DATE STAMPED BELOW

Books not returned on time are subject to a fine of 50c per volume after the third day overdue, increasing to $1.00 per volume after the sixth day. Books not in demand may be renewed if application is made before expiration of loan period.

2m-9,'39 (1176s)

49882